

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 17:08:16 ; Search time 7094.66 Seconds
(without alignments)
10695.346 Million cell updates/sec

Title: US-09-914-651A-2_COPY_14040_16580

Perfect score: 2541

Sequence: 1 agcataattttttttaat.....cttcctctagtgtgaatttt 2541

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estt:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estom:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	245	9.6	486	28	AQ344174
C 2	239	9.4	842	28	AQ745353
C 3	238.8	9.4	538	9	AL041450
C 4	238.6	9.4	609	9	AV760466

C 5	234.6	9.2	1030	14	CD558243
C 6	234.4	9.2	377	9	AI244254
C 7	233.2	9.2	599	9	AL702649
C 8	232.4	9.1	907	14	CF593378
C 9	232.4	9.1	1073	28	BZ610713
C 10	231.8	9.1	546	28	BZ597995
C 11	231.6	9.1	424	9	AA604843
C 12	231.6	9.1	429	28	AQ090219
C 13	231.6	9.1	703	29	AG141722
C 14	231	9.1	530	14	CA422533
C 15	230.6	9.1	547	10	AW504485
C 16	230.2	9.1	533	28	AQ037324
C 17	230	9.1	704	9	AL596482
C 18	229.8	9.0	339	10	AW270258
C 19	229.8	9.0	454	28	BZ892950
C 20	229.8	9.0	702	13	BX645944
C 21	229.6	9.0	803	28	BZ601931
C 22	229.6	9.0	2177	28	AQ839825
C 23	229	9.0	452	9	AL041894
C 24	229	9.0	583	28	AQ544177
C 25	229	9.0	1274	11	BC039325
C 26	228.8	9.0	708	28	AQ530231
C 27	228.4	9.0	759	10	AW976010
C 28	228.4	9.0	800	9	AV755512
C 29	228.2	9.0	450	28	AQ409292
C 30	228	9.0	383	10	BE139267
C 31	228	9.0	657	28	AQ782480
C 32	227.8	9.0	698	28	BZ601203
C 33	227.6	9.0	445	28	AQ349141
C 34	227.6	9.0	735	9	AL042756
C 35	227.4	8.9	487	9	AI859834
C 36	227.4	8.9	509	12	BM511372
C 37	227.4	8.9	556	12	BM510544
C 38	227.4	8.9	656	14	CB115054
C 39	227.4	8.9	771	13	BU617429
C 40	227.4	8.9	775	14	CA428724
C 41	227.2	8.9	354	28	AQ101564
C 42	227.2	8.9	465	9	AI460050
C 43	227.2	8.9	1029	28	BZ610725
C 44	227	8.9	723	13	BX642407
C 45	226.8	8.9	410	13	BX484725

ALIGNMENTS

RESULT 1
AQ344174

LOCUS

DEFINITION

Genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AQ344174 486 bp DNA linear GSS 07-MAY-1999
RPC111-124P5.TV RPCI-11 Homo sapiens genomic clone RPCI-11-124P5,
Genomic survey sequence.
AQ344174 GI:4169070
GSS.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 486)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building (1997)
Unpublished
Other_GSSs: RPC111-124P5.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC

CD558243 AGENCOURT
AI244254 QV90Q05.X
AL702649 DKFZP6860
CF593378 AGENCOURT
BZ610713 WHAC155TR
BZ597995 WHAC155TR
AA604843 ns92f05.s
AQ090219 HS 3009.A
AG141722 Pan trogl
CA422533 UI-H-FL0-
AW504485 UI-HF-BN0
AQ037324 CIT-HSP-2
AL596482 DKFZP761A
AW270258 xp43e01.x
BZ892950 UP 508-19
BX645944 DKFZP781F
BZ601931 WHADO92TR
AQ839825 260L13-C3
AL041894 DKFZP434I
AQ544177 RPCI-11-3
BC039325 Homo sapi
AQ530231 RPCI-11-3
AW976010 EST388119
AV755512 AV755512
AQ409292 HS 5067.A
BE139267 xt58a06.x
AQ782480 HS 3174.A
BZ601203 WHAAMI9TR
AQ349141 RPCI11-11
AL042756 DKFZP434C
AI859834 wm21a08.x
BM511372 ij47b02.Y
BM510544 ij47b02.X
CB115054 K-EST0158
BU617429 UI-H-DF0-
CA428724 UI-H-DF0-
AQ101564 HS 2235.B
AI460050 ap34c01.X
BZ610725 WHACKO1TF
BX642407 DKFZP686C
BX484725 DKFZP686D

library availability, please contact Pieter de Jong (pieter@jmg.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES

Location/Qualifiers
1..486
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7547596"
/db_xref="taxon:9606"
/clone="RPC11-11-124P5"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPC1-11"
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; RPC111 Human Male BAC Library"

ORIGIN

Query Match 9.6%; Score 245; DB 28; Length 486;
Best Local Similarity 87.1%; Pred. No. 2.8e-27;
Matches 269; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 964 TTTTGTGTTGTTTGTGACGAGTCTTACTCTGTCAACCAGGCTGGAGTGTAAATGCATG 1023
DB 8 TTTCCCTTTTGTGAGACAGAGTTTCACTCTGTCAACCAGGCTAGAGTGCATGCATG 67
QY 1024 GTCTCAGCTCAGTGAACCTCGCCCTCCGGGTTCAAGCATTCTTCCACCTCAGCTCC 1083
DB 68 GTCTCAGCTCAGTGAACCTCGCCCTCCGGGTTCAAGCATTCTTCCACCTCAGCTCC 127
QY 1084 CAAAGTAGCTGGGACTACAGCATGAGCCAGCCGCGCTAATTTTGTATTTTAGTA 1143
DB 128 CAGATAGCTGGGACTACAGCGTGTGCCACACCCGCGCAATTTTGTATTTTAGTA 187
QY 1144 GAGACGTGGTTCCACTATGTGGCCAGGCTGATCTCGAACTCTGACCTTGTAAATCCAC 1203
DB 188 GAGACGGGTTTCACTATGTGGCCGCGTGTCTTGAATCTGACCTCGTATCCACC 247
QY 1204 CGGCTCGGCTGCCAAAGTGTGGGATTACAGCGTGAGCCATTGTGCTTGGCCGATTT 1263
DB 248 TGCTCGGCTGCCAAAGTGTGGGATTACAGCGTGAGCCACCAGCGCTGGCCGATGAT 307
QY 1264 TTAATAAAT 1272
DB 308 TTCTATAAT 316

RESULT 2

AQ745353/c
LOCUS
DEFINITION
HS_2276_A1_C05_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2276 Col=9 Row=E, genomic survey
ACCESSION
AQ745353
VERSION
AQ745353.1 GI:5522875
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 842)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
AUTHORS
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE
99380589
PUBMED
10449764
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 2276 row: E column: 9
Seq primer: T7
Class: BAC ends
High quality sequence stop: 842.

FEATURES

Location/Qualifiers
1..842
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2276 Col=9 Row=E"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelosAC11; BAC Clones in E-Coli DH10B"

ORIGIN

Query Match 9.4%; Score 239; DB 28; Length 842;
Best Local Similarity 83.2%; Pred. No. 1.7e-26;
Matches 272; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 964 TTTTGTGTTTGTGACGAGTCTTACTCTGTCAACCAGGCTGGAGTGTAAATGCATG 1023
DB 513 TTTTCATTTTGTGAGACAGAGTCTCACTCTGCTCAGCTGGAGTGCATGCATG 454
QY 1024 GTCTCAGCTCAGTGAACCTCGCCCTCCGGGTTCAAGCATTCTTCCACCTCAGCTCC 1083
DB 453 GTTTCAGCTCAGTGAACCTCGCCCTCCGGGTTCAAGTGGTTCTCCACCTCAGCTCC 394
QY 1084 CAAAGTAGCTGGGACTACAGCATGAGCCAGCCGCGCTAATTTTGTATTTTAGTA 1143
DB 393 CAGATAGCTGGGACTACAGCGTGTGCCACACCCGCGCAATTTTGTATTTTAGTA 334
QY 1144 GAGACGTGGTTCCACTATGTGGCCAGGCTGATCTCGAACTCTGACCTTGTAAATCCAC 1203
DB 333 GAGATGGATTTCATATGTGGCCAGGCTGGTCTTGAATCTGACCTCGTATCCACC 274
QY 1204 CGCCTCGGCTGCCAAAGTGTGGGATTACAGCGTGAGCCATTGTGCTTGGCCGATTT 1263
DB 273 TGCTCGGCTGCCAAAGTGTGGGATTACAGCGTGAGCCATTGTGCTTGGTCTACAGT 214
QY 1264 TTAATAAATGATTTCTTATGTCAGTTT 1290
DB 213 TTGTGTTTAAATCTTCACGTTTTTTT 187

RESULT 3

AL041450
LOCUS
DEFINITION
DKFZp434B117_r1_434 (synonym: htes3) Homo sapiens cDNA clone sequence.
ACCESSION
AL041450
VERSION
AL041450.1 GI:5420801
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 538)
Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and Wienand, S.
AUTHORS
EST (Bloeker, et al.)
TITLE
Unpublished (1999)
JOURNAL
Contact: MIPS
COMMENT
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.

s1 sequence also available.
 This clone (DKFZp434E1117) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Location/Qualifiers
 1..538
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp434E1117"
 /tissue type="testis"
 /dev stage="adult"
 /lab host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /note="vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 9.4%; Score 238.8; DB 9; Length 538;
 Best Local Similarity 89.0%; Pred. No. 2.2e-26;
 Matches 258; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 QY 967 TTTTGTGTTTGGACGAGTCTTACTCTGTCAACCCAGGCTGGAGTGAATGGCATGGTC 1026
 Db 6 TTTTGTGTTTGGACGAGTCTTACTCTGTCAACCCAGGCTGGAGTGAATGGCATGGTC 65
 QY 1027 TCAGCTCACTGCACCTCCGCTCCCGGTTCAAGCAATCTTCCACTCAGCCTCCCAA 1086
 Db 66 TCCACTCACTGCACCTCCGCTCCCGGTTCAAGCAATCTTCCACTCAGCCTCCCGA 125
 QY 1087 GTAGCTGGGACTACAGGATGAGCCAGCCAGCGGCTGAATCTGAACTCTGAACTTGAATAGAG 1146
 Db 126 GTAGCTGGGACTACAGGATGAGCCAGCCAGCGGCTGAATCTGAACTTGAATAGAG 185
 QY 1147 ACCTGGTTCACATATGTTGGCCAGGCTGATCTGAACTCTGAACTTGAATAGAG 1206
 Db 186 ATGGGGTTTCACTATGTTGCCAGGCTGGTCTCAACTCTGAACTTGAATAGAG 245
 QY 1207 CTGGCTGCTCAAGTCTGGATTAAGCGGTGAGCCATGTCCTGGC 1256
 Db 246 CTGGCTGCTCAAGTCTGGATTAAGCGGTGAGCCATGTCCTGGC 295

RESULT 4

AV760466/c 609 bp mRNA linear EST 19-OCT-2000
 LOCUS AV760466 MDS Homo sapiens cDNA clone MDSBKFL1 5', mRNA sequence.
 DEFINITION AV760466.1 GI:10918314
 ACCESSION EST.
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 609)
 Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
 Zeng,L., Xu,S., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z.,
 Lu,G., Yang,Y., Gao,G., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
 Homo sapiens cDNA MDS clones
 Unpublished (2000)
 Contact: Zenguan Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

This clone is available at CHGC in Shanghai.

FEATURES
 Location/Qualifiers
 1..609
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MDSBKFL1"
 /tissue type="Bone marrow"
 /cell type="CD34+ hematopoietic stem/progenitor cell"
 /lab host="BM25.8"
 /clone_lib="MDS"
 /note="vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB"

ORIGIN

Query Match 9.4%; Score 238.6; DB 9; Length 609;
 Best Local Similarity 64.8%; Pred. No. 2.3e-26;
 Matches 381; Conservative 0; Mismatches 204; Indels 3; Gaps 2;
 QY 671 AAAATGAGTTGGGTTTTTAATAATTTCTGAAGTAGGTTTATTGCAATTAATTTT 730
 Db 589 AAATTTATTACTTCATATAATTTTATTTAAACCCCTTTATAAAAATTTTATT 530
 QY 731 TCCTTTAACTTTCAAACTCAAGAAACAGTTGGCCTTGACTCTCTGTGTGGGAAAT 790
 Db 529 TTTTNT 470
 QY 791 TTAACACTACTGGTTTAAATTTCTTTATTTGGTTGTAATATGACTATTTTACGTCA 850
 Db 469 TTTTNT 410
 QY 851 ATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAAATTTTATGTCATAG 910
 Db 409 TTTTNT 350
 QY 911 CAATTTTATGCTTGATATATGACTTTATGTTATATGCTATACAACTAGATTTT 970
 Db 349 CAAACATTAACCAACGAAATAATG--TGATGAAATTTATTTATTTATTTATTTAT 292
 QY 971 GTTGTGTTTTGACCGAGTCTTACTCTGTCACCCAGGCTGGAGTGTATGGCATGTCT 1030
 Db 291 TTATTTTAAAGCAGAGTCTGGCTCTGTGCGCCAGGCTGGAGTGCAGTGGCAGCAT 232
 QY 1031 CTCACCTGCAACCTCCGCTCCCGGTTCAAGCCATTTCTTCCACTCAGCTCCCAAGTAG 1090
 Db 231 CTCACCTGCAAGCTCTGCTCCCTGGTTTATGCACTCTCTTGCTCAGCTCCGAGTAG 172
 QY 1091 CTGGACTACAGCATGAGCCACCGCAGCCGCTAA--TTTTGTTATTTTATGAGAGC 1149
 Db 171 CTGGACTACAGCGCGCCGCCACACACCGCGCTAAATTTTGTATTTTATGAGAGC 112
 QY 1150 TGGTTCACCTATGTTGCCAGGCTGATCTCGAACTCTGAACTTGTAAATCCACCGCTC 1209
 Db 111 GGGTTTACCATGTTAGCCAGGATGCTCGATCTCTGACCTCTGGATCCACCGCCGCC 52
 QY 1210 GGCTGCCCAAAGTCTGGGATTACAGCGGTGAGCCATTGTCCTGGCC 1257
 Db 51 GGCTGCCCAAAGTCTAGGATTACAGCGGTGAGCCACCGCGCCGCC 4

RESULT 5

CD558243/c 1030 bp mRNA linear EST 11-JUN-2003
 LOCUS CD558243 AGENCOURT.1435567 NIH MGC_181 Homo sapiens cDNA clone
 DEFINITION IMAGE:30395963 5', mRNA sequence.
 ACCESSION CD558243
 VERSION CD558243.1 GI:31584311
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1030)
 NIH-MGC http://mgc.nci.nih.gov/.


```
QY 1047 CTTCCCGGTTCAAGCCATTTCTTCCACCTCAGCTCCCAAGTAGCTGGAGCTACAGGCAT 1106
Db 190 CTTCCCGGTTTCAAGCGATTTCTCCACCTCAGCTCCCAAGTAGCTGGAGCTACAGGCAT 249
QY 1107 GAGCCACCGCAGCCGCTAATTTTCTATTTTATTTAGTAGAGAGCTGGTCCACATCTTGG 1166
Db 250 CGGCACACCGCCAGCTAATTTTGTATTTTATTTAGTAGAGAGCGGGTTTCCACATCTTGG 309
QY 1167 CCAGGCTGATCTCGAACTCCTGACCTTGTAACTCCACCCGCTCGGCTGCCAAAGTCTG 1226
Db 310 CCAGGCTGCTCAAACTCCTGACCTTGTATCCACCTCGCTCGGCTGCCAAAGTCTG 369
QY 1227 GAATACAGCGGTAGCCATTTGCTGCTGCGCCGATTTTTTAAAAATGTAATCTTATGTC 1286
Db 370 GAGTACAGGCATAGCCACTGTGCTGCGCGG-GATCTTTGTTTTTTTAACTCTGATAAT 428
QY 1287 GTTTTCATAAGTTTATTTAAATGCAATTTTCCATTTGATGTAAGCTTTCAAATTTATAG 1346
Db 429 TCTGCTTTTGAATCCAAATTTTAAATGCAATTTTCAATTTGATGTAATTTGATTTGAT 488
QY 1347 TATAGTTGTTCTAGTATTTCTTATCTTTTGTAACTCTGTCAGCGCTCTGATAGTGGC 1406
Db 489 TAAGTATGCCATCTTATTTTGTGTTTCTTATGCTGCTCTTTTCTTTCTCTGCTGCC 548
QY 1407 TCTTTT 1412
Db 549 TCTTTT 554
```

RESULT 9

```
BZ610713/c
LOCUS BZ610713 1073 bp DNA linear GSS 08-JUN-2003
DEFINITION WHAC155TR Human MCF7 breast cancer cell line library (MCF7.1) Homo
sapiens genomic clone MCF7_1-15113, genomic survey sequence.
ACCESSION BZ610713
VERSION BZ610713.1 GI:31519274
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1073)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
Gray,J.W. and Collins,C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
22709111
12788976
COMMENT Contact: Volik SV
Colin Collins, lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
```

FEATURES

```
Source
1..1073
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-15113"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library
(MCF7.1)"
/note="Vector: pECBAC1; Site:1: HindIII; This library was
constructed from MCF7 breast cancer cell line by Amplicon
Express (http://www.genomex.com) using their standard
procedure."
```

ORIGIN

```
Query Match 9.1%; Score 232.4; DB 28; Length 1073;
Best Local Similarity 79.5%; Pred. No. 1.5e-25;
Matches 275; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 939 ATTGTTATATGCTATACAACTAGATTTTTTTTGTGTTTTCACCGAGTCTTACTCTGTC 998
Db 475 ATGTCTATTCAGTATTTTGGCATTTTTTTTTTTTTTTTGTGAGACAGAGTCTGCTCTGTC 416
QY 999 ACCCAGGCTGAGTGTAAATGGCATGTCTCAGCTCACTGCAACCTCCGCTCCCGGGTTC 1058
Db 415 GCCCAGGCTGAGTGCAGTGGCATGATCTCAGCTCACTGCAACCTGCTGCCTCACAGTTC 356
QY 1059 AAGCCATCTTCCACTCAGCTCCCAAGTACTGGACTCAGGATGAGCACCAGCAC 1118
Db 355 AAGTGAATCTCTGCTCAGCTCCCAAGTAGTATAGGATTCAGGCAATCGCCACCATGC 296
QY 1119 CCGGCTAATTTTTTGTATTTTATAGTAGAGAGCTGTTTCCACTATGTTGGCAGGCTCATCT 1178
Db 295 CTGGCTAATTTTGTATTTTATAGTAGAGAGGAGTTTACCATGTTGGTCAGGCTGCTCT 236
QY 1179 CGAATCTCTGACCTTTGTAATCCACCCGCTCGGCTGCCAAAGTGTGGATTACAGGCG 1238
Db 235 CGAATCTCTGACCTTTGTGATCCACCTGCTGCGCTCCCAAGTGTGGATTACAGGCA 176
QY 1239 TGAGCCATTTGCTGCGCGGATTTTTTAAATGTAATCTTATGT 1284
Db 175 TGAGCCACTGCTCCCGGCTGCCAATTTTTTAAATCAGATTTTGT 130
```

RESULT 10

```
BZ597995
LOCUS BZ597995 546 bp DNA linear GSS 08-JUN-2003
DEFINITION WHACM46TR Human MCF7 breast cancer cell line library (MCF7.1) Homo
sapiens genomic clone MCF7_1-16G19, genomic survey sequence.
ACCESSION BZ597995
VERSION BZ597995.1 GI:31506457
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 546)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
Gray,J.W. and Collins,C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
22709111
12788976
COMMENT Contact: Volik SV
Colin Collins, lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
```

FEATURES

```
Source
1..546
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-16G19"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library
(MCF7.1)"
/note="Vector: pECBAC1; Site:1: HindIII; This library was
constructed from MCF7 breast cancer cell line by Amplicon
Express (http://www.genomex.com) using their standard
procedure."
```


DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 The following repetitive elements were found in this cDNA
 sequence: 1-57, >(TAAA)n#Simple repeat (matched complement)
 62-347, >ALU (matched complement)
 Seq primer: M13 FORWARD
 POLYA=Yes

FEATURES

Location/Qualifiers

1..530
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FLO-bdg-d-16-0-UI"
 /issue_type="Cell lines"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP FLO"
 /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
 (Pharmacia) with a modified polylinker; Site 1: EcoR I;
 Site 2: Not I; NCI CGAP FLO is a cDNA library derived from
 a pool of mRNA obtained from 4 cell lines from grade III
 chondrosarcoma tissues. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT7T3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is GAGGTGCGTG. The cell line
 was provided by Dr James Martin from University of Iowa.
 TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
 TAG LIB=UI-H-FLO
 TAG_SEQ=GAGGTGCGTG"

ORIGIN

Query Match 9.1%; Score 231; DB 14; Length 530;
 Best Local Similarity 85.1%; Pred. No. 3.3e-25;
 Matches 258; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

964 TTTTGTGTTTGTACCGAGTCTTACTCTGTCAACCGCTGGAGTGAATGGCATG 1023
 Db TTTTGTGTTTGTACCGAGTCTTACTCTGTCAACCGCTGGAGTGAATGGCATG 1023

54 TATTATTATTTTAAACACAGTGTCTGTCTGTCAACCGCTGGAGTGAATGGCATG 113
 Db TATTATTATTTTAAACACAGTGTCTGTCTGTCAACCGCTGGAGTGAATGGCATG 113

1024 GTCTCAGTCACTGCAACCTCCGCTCCCGGTTCAAGCCATTCTCCACTGAGCTCC 1083
 Db GTCTCAGTCACTGCAACCTCCGCTCCCGGTTCAAGCCATTCTCCACTGAGCTCC 1083

114 ATCTTGCTCACTGCAACCTCCGCTCCCGGTTCAAGCCATTCTCCACTGAGCTCC 173
 Db ATCTTGCTCACTGCAACCTCCGCTCCCGGTTCAAGCCATTCTCCACTGAGCTCC 173

1084 CAAGTAGCTGGACTACAGGATGAGCCACCGCCGGCTAAATTTTGTATTTTAGTA 1143
 Db CAAGTAGCTGGACTACAGGATGAGCCACCGCCGGCTAAATTTTGTATTTTAGTA 1143

174 CAAGCAGCTGGACTACAGGATGAGCCACCGCCGGCTAAATTTTGTATTTTAGTA 233
 Db CAAGCAGCTGGACTACAGGATGAGCCACCGCCGGCTAAATTTTGTATTTTAGTA 233

1144 GAGACGTGGTTCCACTATGTGGCAGGTGATCTCGAACTCTGACCTTGTATTCACC 1203
 Db GAGACGTGGTTCCACTATGTGGCAGGTGATCTCGAACTCTGACCTTGTATTCACC 1203

234 GAGACGTGGTTCCACTATGTGGCAGGTGATCTCGAACTCTGACCTTGTATTCACC 293
 Db GAGACGTGGTTCCACTATGTGGCAGGTGATCTCGAACTCTGACCTTGTATTCACC 293

1204 CGCTCGGCTGCGCAAGTGTGGATTACAGGCTGAGCCATTGTGCTTGGCGGATTT 1263
 Db CGCTCGGCTGCGCAAGTGTGGATTACAGGCTGAGCCATTGTGCTTGGCGGATTT 1263

294 CACTCGGCTGCGCAAGTGTGGATTACAGGCTGAGCCATTGTGCTTGGCGGATTT 353
 Db CACTCGGCTGCGCAAGTGTGGATTACAGGCTGAGCCATTGTGCTTGGCGGATTT 353

1264 TTA 1266
 Db TTA 1266

354 TGA 356
 Db TGA 356

RESULT 15
 AW504485
 LOCUS
 DEFINITION
 UI-HF-BNO-alc-h-01-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone
 IMAGE:3080665 5', mRNA sequence.
 ACCESSION
 AW504485

VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AW504485.1 GI:7142152

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 547)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Eco RI site shown at the beginning of the sequence.

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: M.B. Soares Lab

cDNA Library Arrayed by: M.B. Soares Lab

DNA Sequencing by: M.B. Soares Lab

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/brp/image/image.html

The following repetitive elements were found in this cDNA

sequence:

25-105, >ALU 118-139, >(TA)n#Simple_repeat 159-445, >ALU

Seq primer: M13 Forward.

Location/Qualifiers

1..547

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3080665"

/tissue_type="lymph"

/cell_type="germinal center B cells"

/cell_line="MGC85"

/lab_host="DH10B (LTI)"

/clone_lib="NIH_MGC_50"

/note="Vector: pT7T3-Pac; Site 1: NotI; Site 2: Eco RI;

Constructed from size fractionated cytoplasmic mRNA

(3.5-4.4kb). Directionally cloned. Cells provided by

Louis M. Staudt, Ph.D. Library preparation by Maria de

Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

Query Match 9.1%; Score 230.6; DB 10; Length 547;

Best Local Similarity 75.3%; Pred. No. 3.8e-25;

Matches 287; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

919 ATTGCTTGATATGATGATTTATTTATATGGCTATACAACTAGATTTTGTGTTT 978

Db ATTGCTTGATATGATGATTTATTTATATGGCTATACAACTAGATTTTGTGTTT 978

108 AGTGTGTGAAT 167

Db AGTGTGTGAAT 167

979 TGACCGAGTCTTACTCTGTCAACCGCTGGAGTGAATGGCATGTTCTCAGCTCCTGC 1038

Db TGACCGAGTCTTACTCTGTCAACCGCTGGAGTGAATGGCATGTTCTCAGCTCCTGC 1038

168 AGCGGAGTCTCACCCTGTTGTCCAGCTGGAGTGAATGGCATGTTCTCAGCTCCTGC 227

Db AGCGGAGTCTCACCCTGTTGTCCAGCTGGAGTGAATGGCATGTTCTCAGCTCCTGC 227

1039 AACCTCGGCTCCCGGTTCAAGCCATTCTTCACCTCAGCCCTCCCAAGTAGTGGGACT 1098

Db AACCTCGGCTCCCGGTTCAAGCCATTCTTCACCTCAGCCCTCCCAAGTAGTGGGACT 1098

228 AACCTCGGCTCCCGGTTCAAGCCATTCTTCACCTCAGCCCTCCCAAGTAGTGGGACT 287

Db AACCTCGGCTCCCGGTTCAAGCCATTCTTCACCTCAGCCCTCCCAAGTAGTGGGACT 287

1099 ACAGGATGAGCCACCGCCAGCTAAATTTTGTATTTTAGTAGAGAGCTGGTCCAC 1158

Db ACAGGATGAGCCACCGCCAGCTAAATTTTGTATTTTAGTAGAGAGCTGGTCCAC 1158

288 ACAGGCGCCCGCCACCGCCAGCTAAATTTTGTATTTTAGTAGAGAGCTGGTCCAC 347

Db ACAGGCGCCCGCCACCGCCAGCTAAATTTTGTATTTTAGTAGAGAGCTGGTCCAC 347

1159 TATGTTGGCCAGGCTGATCTCGAAGTCTGACCTTGAATCCACCGCTCGGCTGGCCA 1218

Db TATGTTGGCCAGGCTGATCTCGAAGTCTGACCTTGAATCCACCGCTCGGCTGGCCA 1218

348 CATGTTGGCCAGGCTGGTCTTGAACCTCTGACCTTGAATCCACCGCTCGGCTGGCCA 407

Db CATGTTGGCCAGGCTGGTCTTGAACCTCTGACCTTGAATCCACCGCTCGGCTGGCCA 407

1219 AAGTGTGGGATTACAGGCTGAGCCATTGTGCTCGCGGATTTTAAAGATGATTTC 1278

Db AAGTGTGGGATTACAGGCTGAGCCATTGTGCTCGCGGATTTTAAAGATGATTTC 1278

408 AAGTGTGGGATTACAGGCTGAGCCATTGTGCTCGCGGATTTTAAAGATGATTTC 1299

Db AAGTGTGGGATTACAGGCTGAGCCATTGTGCTCGCGGATTTTAAAGATGATTTC 1299

1279 TTAGTGTGATTTTCAATGATT 1299

Db TTAGTGTGATTTTCAATGATT 1299

468 TGAAGTCAATTTGAATGGGT 488

Db TGAAGTCAATTTGAATGGGT 488

us-09-914-651a-2_copy_14040_16580.rst

Mon May 24 08:12:35 2004

Search completed: May 21, 2004, 23:03:24
Job time : 7099.66 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 20:49:36 ; Search time 1146.81 Seconds
(without alignments)

10069.828 Million cell updates/sec

Title: US-09-914-651A-2_COPY_14040_16580

Perfect score: 2541

Sequence: 1 agcataatttttttttaatt.....cttctctagtgtgaatttt 2541

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	312	12.3	335	16	US-10-431-791-3
2	271.2	10.7	174566	13	US-10-235-192A-37
3	271.2	10.7	174566	15	US-10-020-141-1
4	262	10.3	454	16	US-10-431-791-1
5	252	9.9	157875	10	US-09-935-464-1
6	252	9.9	157875	15	US-10-125-835-1
7	245	9.7	257	16	US-10-431-791-18
8	241.4	9.5	246144	16	US-10-085-117-226
9	239.8	9.4	21470	9	US-09-764-847-1157
10	239.8	9.4	21470	15	US-10-092-154-1157
11	239.6	9.4	133632	13	US-10-087-192-1810
12	239.4	9.4	16086	9	US-09-764-877-2385
13	239.4	9.4	16086	16	US-10-242-515-2385
14	239	9.4	750	13	US-10-027-632-284167

c 15 239 9.4 750 16 US-10-027-632-284167
c 16 238.8 9.4 538 10 US-09-764-891-7553
c 17 238.8 9.4 538 10 US-09-764-891-7555
c 18 238 9.4 15266 9 US-09-764-877-3797
c 19 238 9.4 15266 16 US-10-242-515-3797
c 20 238 9.4 15271 9 US-09-764-877-3798
c 21 238 9.4 15271 16 US-10-242-515-3798
c 22 238 9.4 143068 9 US-09-967-768A-316
c 23 237.8 9.4 126512 9 US-09-804-474A-3
c 24 236.4 9.3 62804 14 US-10-036-960-3
c 25 235.4 9.3 1877 13 US-10-027-632-256674
c 26 235.4 9.3 1877 16 US-10-027-632-256674
c 27 235.2 9.3 656 13 US-10-027-632-226967
c 28 235.2 9.3 656 16 US-10-027-632-226967
c 29 235 9.2 797 13 US-10-027-632-127789
c 30 235 9.2 797 16 US-10-027-632-127789
c 31 234.8 9.2 650 13 US-10-027-632-214024
c 32 234.8 9.2 650 16 US-10-027-632-214024
c 33 234.8 9.2 13444 9 US-09-764-877-3660
c 34 234.8 9.2 13444 16 US-10-242-515-3660
c 35 234.6 9.2 66973 13 US-10-087-192-574
c 36 234.6 9.2 176001 17 US-10-210-556-27
c 37 234.6 9.2 186739 17 US-10-210-556-19
c 38 234.4 9.2 56423 16 US-10-292-798-181
c 39 234 9.2 539 13 US-10-027-632-15927
c 40 234 9.2 539 16 US-10-027-632-15927
c 41 233.6 9.2 1741 13 US-10-027-632-253802
c 42 233.6 9.2 1741 16 US-10-027-632-253802
c 43 233.6 9.2 40491 13 US-10-087-192-1426
c 44 233.4 9.2 656 13 US-10-027-632-27832
c 45 233.4 9.2 656 16 US-10-027-632-27832

ALIGNMENTS

RESULT 1
US-10-431-791-3
; Sequence 3, Application US/10431791
; Publication No. US20030235874A1
; GENERAL INFORMATION:
; APPLICANT: Kao, Chinghai
; APPLICANT: Lee, Sang-Jin
; APPLICANT: Kim, Hong-Sup
; APPLICANT: Lee, KangRyul
; APPLICANT: Yu, Rong
; TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and
; FILE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 1857-ARTI.0222US
; CURRENT APPLICATION NUMBER: US/10431,791
; CURRENT FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: 60/378,920
; PRIOR FILING DATE: 2002-05-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-10-431-791-3

Query Match 12.3%; Score 312; DB 16; Length 335;
Best Local Similarity 99.4%; Pred. No. 1e-44;
Matches 334; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 721 AATATTTTTCCTTAACTTTCAACTCAAGAAACAGTTGGCTTGACTCTGTTT 780
|||||
Db 1 AATATTTTTCCTTAACTTTCAACTCAAGAAACAGTTGGCTTGACTCTGTTT 60
|||||
QY 781 GTGGAATAATTTAACTACTGTTTAACTTTTATTGTTTGAATATGACTATTATACG 840
|||||

Db	61	GTGGAAAAATTTTAAACACTACTGGTTTAAATTTCTTTATTTGGTTGTAATATGACTATTTTTCAG	120
Qy	841	TCATATAACAATTTTATTTATTTGTTTAAATGACTTTATTT - GTTTGTCATATGATAATTTT	899
Db	121	TCATATAACAATTTTATTTGTTTAAAGACTTTATTTGGTTTGTGCATATGATAATTTT	180
Qy	900	ATGTCATAGAACAAATTTTATTTGCTTTGATATATGACTTTATTTGTTATATGCGTATACAAC	959
Db	181	ATGTCATAGAACAAATTTTATTTGCTTTGATATATGACTTTATTTGTTAT - TGGCTATACAAC	239
Qy	960	TAGATTTTATTTGTTTGGTTTACCGAGTCTTACTCTGTCCACCCAGGCTGGAGTGAATGG	1019
Db	240	TAGATTTTATTTGTTTGGTTTACCGAGTCTTACTCTGTCCACCCAGGCTGGAGTGAATGG	299
Qy	1020	CATGGTCTAGCTCACTGCAACCTCCGCTCCCGGG	1055
Db	300	CATGGTCTAGCTCACTGCAACCTCCGCTCCCGGG	335
RESULT 2			
US-10-235-192A-37/c			
; Sequence 37, Application US/10235192A			
; Publication No. US20040043389A1			
; GENERAL INFORMATION:			
; APPLICANT: McCarthy, Jeanette			
; TITLE OF INVENTION: Methods and Compositions for Identifying			
; TITLE OF INVENTION: Risk Factors for Abnormal Lipid Levels and the Diseases			
; TITLE OF INVENTION: and Disorders Associated Therewith			
; FILE REFERENCE: MMI-011			
; CURRENT APPLICATION NUMBER: US/10/235,192A			
; CURRENT FILING DATE: 2002-09-04			
; NUMBER OF SEQ ID NOS: 49			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 37			
; LENGTH: 174566			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-235-192A-37			
Query Match 10.7%; Score 271.2; DB 13; Length 174566;			
Best Local Similarity 53.0%; Pred. No. 1.6e-36;			
Matches 802; Conservative 0; Mismatches 678; Indels 34; Gaps 9;			
Qy	880	GTTTGTGCATATGATAATTTTATGTCATAGAACAAATTTTATGCTTGATATATGACTTTA	939
Db	37228	GTTTGTGCATGATAGCGCTTTATTTATTTGAGGTATGTTTCCCTTTATAGTCAATCAGATTG	37169
Qy	940	TGTTATATGCTATA - CAACATAGATTTTGTGTTTGTGTTTGCACCGAGTCTTACTCTGTC	998
Db	37168	TATATTTATTTATTTATTTATTTTATTTTATTTTGTGAGTGGAGTCTGCTGTGTC	37109
Qy	999	ACCAGGCTGGAGTGTAAATGATGCTCAGCTCAGTCACTGCAACCTCCGCTCCGGGTTTC	1058
Db	37108	ATGCAGGCTGGAGTGCAGTGACACAAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTC	37049
Qy	1059	AAGCAATCTTCCACCTCAGCTCCAGTAGCTGGGACTACAGGCATGAGCCACGCGAC	1118
Db	37048	AAGGGAATCTCTCGCTCAGCTCCCAAGTAACCTGGTACTACAGGCATGACCAACATGC	36989
Qy	1119	CGGCT - AATTTTGTATTTTATAGACAGTGGTCCACTATGTTGGCCAGGCTGATC	1177
Db	36988	CTGGCTAAATTTTGTATTTTATAGATGGGGTTTACCATGTTGGCCAGGCTGATG	36929
Qy	1178	TCGAATCCTGACCT - TGTAAATCCACCGGCTCGGCTGCGCAAGTCTGGGATTCAG	1235
Db	36928	TCGAATCCTGACCTCAAGTGATCGGCTGCTGGTCTCCCAAGTGTGGAAATTCAG	36869
Qy	1236	GGTGAGGCATGTGCTGGCGGATTTTAAAAAATGATTTCTTATGTCAGTTTTCATA	1295
Db	36868	GTGTGAGGCATGTGACCCAGCCAGATTTTATTTGTTAGAAATCAAGGATATGTAATGT	36809
Qy	1296	AGTTTATTTAAATGCAATTTTCCATTTGATGTAAGCTTTTCAAAATTTATAGTATAGTTGT	1355

APPLICANT: Lee, KangRyul
APPLICANT: Yu, Rong
TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and
TITLE OF INVENTION: Methods of Use Thereof
FILE REFERENCE: 1857-ARTI.0222US
CURRENT APPLICATION NUMBER: US/10/431,791
CURRENT FILING DATE: 2003-05-08
PRIOR APPLICATION NUMBER: 60/378,920
PRIOR FILING DATE: 2002-05-08
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 454
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Sequence
US-10-431-791-1
Query Match 10.3%; Score 262; DB 16; Length 454;
Best Local Similarity 100.0%; Pred. No. 5.6e-36;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 721 AATTATTTTCTTTAACTTTCAAACTCAAGGAAACAGTTGGCCTTGACTCTGTTT 780
Db 193 AATTATTTTCTTTAACTTTCAAACTCAAGGAAACAGTTGGCCTTGACTCTGTTT 252
QY 781 GTGGAAATTTTAACTACTGTTTAAATTTCTTTATTTGTTGTAATGACTATTTTACG 840
Db 253 GTGGAAATTTTAACTACTGTTTAAATTTCTTTATTTGTTGTAATGACTATTTTACG 312
QY 841 TCATATAACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 900
Db 313 TCATATAACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 372
QY 901 TGTCTAGAACAAATTTTATTTGTTGTAATGACTTTATTTGTTGTAATGACTATTTTAC 960
Db 373 TGTCTAGAACAAATTTTATTTGTTGTAATGACTTTATTTGTTGTAATGACTATTTTAC 432
QY 961 AGATTTTGTGTTTTCAC 982
Db 433 AGATTTTGTGTTTTCAC 454
RESULT 5
US-09-935-464-1
Sequence 1, Application US/09935464
Publication No. US20030027153A1
GENERAL INFORMATION:
APPLICANT: Meyer, Joanne
APPLICANT: Barrington-Martin, Rory
APPLICANT: Parker, Alexander
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
FILE REFERENCE: 3322/1H702 US1
CURRENT APPLICATION NUMBER: US/09/935,464
CURRENT FILING DATE: 2001-08-23
PRIOR APPLICATION NUMBER: US 09/757,300
PRIOR FILING DATE: 2001-01-09
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 157875
TYPE: DNA
ORGANISM: Homo sapiens
US-09-935-464-1
Query Match 9.9%; Score 252; DB 10; Length 157875;
Best Local Similarity 54.0%; Pred. No. 3.4e-33;
Matches 862; Conservative 0; Mismatches 635; Indels 99; Gaps 13;
QY 805 TAATTTCTTTATTTGTTGTAATGACTATTTTACGTCATATAACAATTTTATTTGTTG 864

Db 33292 TAGGCTTTCTTTGATGGAGACATTTATTACAGATTTAAATCTCTTTACTCATATTTTG 33351
QY 865 TTAATATGACTTTATTGTTTGTTCATATGATAAATTTTATGTCATAGAACAAATTTTATTTGTT 924
Db 33352 TCTGTTTCAGATTTTCTATTCTTCTTCTGTTTCTATCTTGTAGTTGCAAGATACATATGTC 33411
QY 925 TGATATATGACTTTATTGTTTATATGCTATACAACTAGATTTTATTTGTTTGTGTTTACCG 984
Db 33412 CACGAATTTATCAACTTCTGCTAGGTTTATTTTGTGTTTATTTTGTGTTTGTGATG 33471
QY 985 AGTCTTACTCTGTCACCCAG-----GCTGAGAGTGAATGGCATGCTCTCAGCTCACTGC 1038
Db 33472 AGTCTGACTTTTGTGCCAGGCTGGAGCTGGAGTGCAGTGCCTGATCTCAGCTCACTGC 33531
QY 1039 AACCTCCGCTCCCGGTTCAAGCCATTTCTCCACTCAGCCTCCCAAGTAGCTGGACT 1098
Db 33532 AAGCTGGCTCCAGGTTCCAGCCATTTCTCTGCTCAGCCTCCGAGTAGCTGGACT 33591
QY 1099 --ACAGGATGAGCCACCCAGCTGCTTATTTT-----GTATTTTCTAGTACGCTGG 1152
Db 33592 ACACAGGCTCCGCTCCAGCCATGCTGCTGCTTATTTTGTATTTTATTTAGTAGAGTGGG 33651
QY 1153 TTCCCATGTTTGGCAGGCTGATCTCGAACTCCTGACCTTGTAAATCCACCCGCTCGGC 1212
Db 33652 TTTCTGTTTGGCAGGATGCTCTGATCCCTGACCTCGTATCTGCCGCTCGGC 33711
QY 1213 CTGCCAAAGTGTGGGATTTACAGGCTGAGCCATTTGCTGCTGCGGATTTTAAATAAT 1272
Db 33712 CTGCCAAAGTGTGGGATTTACAGGCTGAGCCACTGCACCCGACGATTTCTGCTAGGTT 33771
QY 1273 GTATCTTATGTCAGTTTTCATAGTTTATTTAAATGCAATTTTCCATTTGATGTAAGC 1332
Db 33772 TTTCTAATTAATTTGGCAAT----- 33789
QY 1333 TTTCAAATTTATAGTATAGTTGTTCTCTAGTATTTTCTTATCTTTTGTAAATCTCTTCAGCG 1392
Db 33790 -----TAGTTGTTCTATAGTAGTCTATGATCTTTGTTATTTGTCATATG 33833
QY 1393 TCTGTAGATGCTCTCTTTTAAATAAATAATTTATTTGTTGCGGTTTTT--GCTATTTT 1450
Db 33834 TAATTTGATGCTCTCTCTTTTATTTGATTTTATTTGATTTTATTTGATCTTTTACCTTTT 33893
QY 1451 TTTTCTTATGCTCTTTGAGAGGATATGTCAAATTTTACTAGTGTATCCAAAGATTAAC 1510
Db 33894 TTTCTTGGTTAATCTTTCTAATGTTTGTCAATTTTGTGTTTATCTTTTCAAAAAAGCACT 33953
QY 1511 TTGGGCTGTCGAATCTTTTCTCATCTATCTTTGCTTTATATTTTATTTATTTCTGTTCTTG 1570
Db 33954 TTTTCTTGTATCTTTTAAATTTTATGACTCTATATTTGTTGATTTCTGCTCTCTATCTTT 34013
QY 1571 TTTTATAATTTGCTCTTTTATCTTTCTTTGTTTATCTTTGCTGTTCTTTGTTGAAAATCCTC 1630
Db 34014 ATTTATTTCTCTCTCTCTTTAAATTTTTC--GGATTTTGGTTTCTCGTTGTTCTAGGCCCTTG 34072
QY 1631 AGTAGAATGCTTTAATTTATGATTTGATTTGATTTCTTTCTTTCTTTCTATCTATGATTTTAGA 1690
Db 34073 AAATGCTTTATTTAGTTTGTTTTATTTGCAATCTCTCTACTTTTATAATATAGAGCCATTTAT 34132
QY 1691 GCCATAAAATTTCCCTTTTAACTTTCCCTTTTCCACTTCAACTTACATCTCACAAAATTTGGATT 1750
Db 34133 GCTATTAATTTTCTTTT-----AGACTACTTTTGTGTTGTTTCTTATGATTTGGTA 34186
QY 1751 AGGAGTAGTTTAAATATCATTTAGTATCTAAATATTTTAAATTTTCTGAT-----TTTCTTC 1807
Db 34187 TATTGTGATCTATTTTCTATTTTCAATTTTCAAGAAATTTTAAATTTCTCTTCTTAAATTTATTA 34246
QY 1808 TTTGATCCTGCAACTATTTTACAGTATTTTAAATCTT-----GAATATAAGATTTGTTATT 1866
Db 34247 ATTTAGCCCAATTTGCTCAATTTAGGAGCATTTGTTTAAATTTTCTGTTTATTTATATGTTTCCA 34306
QY 1867 GTTATTTGTTGATCTGATCTCTAAATTTGAATATATTTAGATCAGATTAATTTGTTGTTGTA 1926
Db 34307 AAGTTGTTCTTGGTATTAATTTTATTTTATTTTCCATGATCAGAAAAGATATTTGATA 34366


```
/ NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1157
; LENGTH: 21470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1157

Query Match
Best Local Similarity 9.4%; Score 239.8; DB 9; Length 21470;
Matches 277; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 964 TTTTCTGCTGCTTTTGGACCGAGCTTACTCTGTACACCCAGGCTGGAGTGAATGGCAG 1023
Db 8153 TTTTCTGCTGCTTTTGGACCGAGCTTACTCTGTACACCCAGGCTGGAGTGAATGGCAG 8212
QY 1024 GTCTCAGCTCACTGCAACTCGGCTCCCGGGTTCAGGCCATCTTCCACCTCAGCCTCC 1083
Db 8213 GTCTCGGCTCACTGCAACTCGGCTCCCGGGTTCAGGCCATCTTCCACCTCAGCCTCC 8272
QY 1084 CAAGTAGCTGGAGCTACAGGCATGAGCCACCGCCGCTAAATTTTGTATTTTAGTA 1143
Db 8273 TGAGTAGCTGGAGCTACAGGCATGAGCCACCGCCGCTAAATTTTGTATTTTAGTA 8332
QY 1144 GAGACGTGTTCCACTATGTTGGCCAGGCTGATCTCGAAGCTCTGACCTGTAATCCACC 1203
Db 8333 GAGACAGGTTTCACTATGTTGGCCAGGCTGATCTCGAAGCTCTGACCTGTAATCCACC 8392
QY 1204 CGCCTCGGCTGCCAAAGTGTGGGATACAGGCTGAGCCATTTGCTGGCCGATTTT 1263
Db 8393 TGCTCAGCCTCCAAAGTGTGGGATACAGGCTGAGCCATTTGCTGGCCGATTTT 8452
QY 1264 TTAATAAATGATTTCTTATGTCAGTTTTCATAAGTTTGA 1302
Db 8453 GTTTATGTGAAGAAGATTTCTCAGTGTTTACATCTGTCA 8491

RESULT 10
US-10-092-154-1157
; Sequence 1157, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC0909C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1157
; LENGTH: 21470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1157

Query Match
Best Local Similarity 9.4%; Score 239.8; DB 15; Length 21470;
Matches 277; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 964 TTTTCTGCTGCTTTTGGACCGAGCTTACTCTGTACACCCAGGCTGGAGTGAATGGCAG 1023
Db 8153 TTTTCTGCTGCTTTTGGACCGAGCTTACTCTGTACACCCAGGCTGGAGTGAATGGCAG 8212
QY 1024 GTCTCAGCTCACTGCAACTCGGCTCCCGGGTTCAGGCCATCTTCCACCTCAGCCTCC 1083
Db 8213 GTCTCGGCTCACTGCAACTCGGCTCCCGGGTTCAGGCCATCTTCCACCTCAGCCTCC 8272
QY 1084 CAAGTAGCTGGAGCTACAGGCATGAGCCACCGCCGCTAAATTTTGTATTTTAGTA 1143
Db 8273 TGAGTAGCTGGAGCTACAGGCATGAGCCACCGCCGCTAAATTTTGTATTTTAGTA 8332
QY 1144 GAGACGTGTTCCACTATGTTGGCCAGGCTGATCTCGAAGCTCTGACCTGTAATCCACC 1203

RESULT 11
US-10-087-192-1810/c
; Sequence 1810, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1810
; LENGTH: 133632
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1810

Query Match
Best Local Similarity 9.4%; Score 239.6; DB 13; Length 133632;
Matches 284; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 909 AACATTTTATTTGCTTGATATATGATTTATTTGTTATATATGCTGATACAACTAGATTTT 968
Db 72842 AGCTATTATTATTTCTAAATATATATAGCATTCATTTTAAAGGTATATATAGGATTTT 72783
QY 969 TTGTTGTTTGTGACCGAGTCTTACTCTGTACACCCAGGCTGGAGTGAATGGCATGCTC 1028
Db 72782 TTTTCTGAGACAGAGTCTTCTGTGCGCCAGGCTGGAGTGAATGGCATGATCTC 72723
QY 1029 AGCTCACTGCAACCTCCGCTCCCGGGTTCAGGCCATTTCTTCCACCTCAGCCTCCCAAGT 1088
Db 72722 GGCACACTGCAACCTCCGCTCCCGGGTTCAGGCCATTTCTTCTGCTCAGCCTCCCTGT 72663
QY 1089 AGCTGGAGCTACAGGCATGAGCCACCGCCGCTAAATTTTGTATTTTAGTAGAGAC 1148
Db 72662 AGCTGGAGTATACAGGCACCGCCGCTAAATTTTGTATTTTAGTAGAGAT 72603
QY 1149 GTCTGTTCCACTATGTTGGCCAGGCTGATCTCGAAGCTCTGACCTGTAATCCACCCGCT 1208
Db 72602 GGGGTTTCACTATGTTGGTTCAGGCTGGTCTGCAAGCTCTGACCTGATGATCCACCT 72543
QY 1209 CGGCTGCCAAAGTGTGGGATACAGGCTGAGCCATTTGCTGCTGCCGATTTTGA 1266
Db 72542 CGGCTGCCAAAGTGTGGGATACAGGCTGAGCCATTTGCTGCTGCCGATTTTGA 72485

RESULT 12
US-09-764-877-2385
; Sequence 2385, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
```

```

; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2385
; LENGTH: 16086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2385

Query Match          9.4%; Score 239.4; DB 9; Length 16086;
Best Local Similarity 70.8%; Pred. No. 2e-31;
Matches 318; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 817 TGGTTGTAATATGACATATTTTACGTCATATAACAATTTTATTTGTTGTTAAATGACTTT 876
Db 13032 TGTAAATGTTTGGCTGTGATAGGCAATTTAGTCTACTGATGCTGATTTATTTATTTATGGAT 13091

QY 877 ATTGTTTGTCAATGATAATTTTATGTCATAGAACAAATTTTATTTGCTTTGATATATGACT 936
Db 13092 ATATTGGGTTTATATCTCCCATCACAATTTTGTGCTATTTGTCCCACTCTTGTTTCT 13151

QY 937 TTATTGTTTATATGGCTATACAACTAGATTTTTTGTGTTTGTGTTGACCGAGTCTTACTCTG 996
Db 13152 TTTTCTCTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 13211

QY 997 TCACCCAGGCTGGAGTGTAATGTCATGCTCAGCTCAGTCACTGCAACCTCCGCGCTCCCGGGT 1056
Db 13212 TCACCCAGGCTGGAGTGCAATGCGTGATCTCGGCTCAGTCACTGCAACCTCCGCGCTCCCGGGT 13271

QY 1057 TCAAGCCATTCTTCCACCTCAGCTCCCAAGTAGCTGGGACTACAGGCATGAGCCACCGC 1116
Db 13272 TCATGCCATCTCTCTGCTCAGCTCCCAAGTAGCTGGGACTACAGGCATGAGCCACCAT 13331

QY 1117 ACCCGCTAATTTTGTATTTTATTTAGTAGAGACGTGTTTCACTATGTTGGCCAGGCTGAT 1176
Db 13332 GGCGGCTAATTTATTTGTTATTTTATTTAGTAGAGACGGGTTTCACTATGTTGGCCAGGCTGAT 13391

QY 1177 CTCGAACCTCTGACCTTGATCCACCGCTCGGCTCGCAAGTCTGGGATTACAGG 1236
Db 13392 CTCGAACCTCTGACCTTGATCTGCGCGCTTGGCTCCCAAGTCTGGGATTACAGG 13451

QY 1237 CGTGAGCCATTGTCCTGGCCGATTTTTT 1265
Db 13452 CGTGAGCCACCGCGCTGGCTTTTCT 13480

```

RESULT 13

```

US-10-242-515-2385
; Sequence 2385, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496

```

RESULT 14

```

US-10-027-632-284167/c
; Sequence 284167, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

```

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284167
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-284167

Query Match          9.4%; Score 239; DB 13; Length 750;
Best Local Similarity 74.6%; Pred. No. 6.7e-32;
Matches 299; Conservative 1; Mismatches 101; Indels 0; Gaps 0;

QY 865 TTAATGACTTTATTGTTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATTGCT 924
Db      |||||
QY 451 TAAATTTTAAAGAAATTTCTTTTGCAAAATATCTTTTCAATATATGTCCTATATGC 392
Db      |||||
QY 925 TGATATATGACTTTTATGTTTATATGCTATATACAACTAGATTTTGTGTTTGTGACCG 984
Db      |||||
QY 391 TTGATGCTTATCATTTGTTTGTCTCTATCTCATTTTATTTTATCTTTTGTGAGACGG 332
Db      |||||
QY 985 AGTCTTACTCTGTCAACCAGGCTGGAGTGTAAATGGCATGGTCTCAGCTCACTGCAACCTC 1044
Db      |||||
QY 331 AATCTCGCTCTGTCAACCAGTTTGCAGTGCAGTGCATGATCTCGGCTCACTGCAACCTC 272
Db      |||||
QY 1045 CGCTCCCGGGTTCAAGCCATTCTTCCACCTCAGCTCCCAAGTAGTGGGACTACAGGC 1104
Db      |||||
QY 271 CGCTCCCGAGTTCAAGCAATCTTCTGCTCAGCTCCCAAGTAGTGGGACTACAGGC 212
Db      |||||
QY 1105 ATGAGCCACCGCACCCGGCTAAATTTTGTATTTTATGATAGAGACGTGTTCCACTATGTT 1164
Db      |||||
QY 211 AAGTGCCACCAATCTGGCTAAATTTTGTATTTTATGATAGAGACAGGGTTTCCACTACT 152
Db      |||||
QY 1165 GGCAGGCTGATCTCGAACTCTGACCTTGTAAATCCACCCGCTCGGCTGCCAAAGTGC 1224
Db      |||||
QY 151 GGCAGGCTGCTTGTAACTCATGACCTCGTGAATCCACCTGCTGCGCTGCCAAAGTGC 92
Db      |||||
QY 1225 TGGATTACAGGCGTGAGCCATTGTGCTGCGCGATTTT 1265
Db      |||||
QY 91 TGGATTACAGGCGTGAGCCACCTGCCCCAGTACTTTTAT 51
Db      |||||

Search completed: May 22, 2004, 02:28:16
Job time : 1154.81 secs
```

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284167
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-284167

Query Match          9.4%; Score 239; DB 13; Length 750;
Best Local Similarity 74.6%; Pred. No. 6.7e-32;
Matches 299; Conservative 1; Mismatches 101; Indels 0; Gaps 0;

QY 865 TTAATGACTTTATTGTTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATTGCT 924
Db      |||||
QY 451 TAAATTTTAAAGAAATTTCTTTTGCAAAATATCTTTTCAATATATGTCCTATATGC 392
Db      |||||
QY 925 TGATATATGACTTTTATGTTTATATGCTATATACAACTAGATTTTGTGTTTGTGACCG 984
Db      |||||
QY 391 TTGATGCTTATCATTTGTTTGTCTCTATCTCATTTTATTTTATCTTTTGTGAGACGG 332
Db      |||||
QY 985 AGTCTTACTCTGTCAACCAGGCTGGAGTGTAAATGGCATGGTCTCAGCTCACTGCAACCTC 1044
Db      |||||
QY 331 AATCTCGCTCTGTCAACCAGTTTGCAGTGCAGTGCATGATCTCGGCTCACTGCAACCTC 272
Db      |||||
QY 1045 CGCTCCCGGGTTCAAGCCATTCTTCCACCTCAGCTCCCAAGTAGTGGGACTACAGGC 1104
Db      |||||
QY 271 CGCTCCCGAGTTCAAGCAATCTTCTGCTCAGCTCCCAAGTAGTGGGACTACAGGC 212
Db      |||||
QY 1105 ATGAGCCACCGCACCCGGCTAAATTTTGTATTTTATGATAGAGACGTGTTCCACTATGTT 1164
Db      |||||
QY 211 AAGTGCCACCAATCTGGCTAAATTTTGTATTTTATGATAGAGACAGGGTTTCCACTACT 152
Db      |||||
QY 1165 GGCAGGCTGATCTCGAACTCTGACCTTGTAAATCCACCCGCTCGGCTGCCAAAGTGC 1224
Db      |||||
QY 151 GGCAGGCTGCTTGTAACTCATGACCTCGTGAATCCACCTGCTGCGCTGCCAAAGTGC 92
Db      |||||
QY 1225 TGGATTACAGGCGTGAGCCATTGTGCTGCGCGATTTT 1265
Db      |||||
QY 91 TGGATTACAGGCGTGAGCCACCTGCCCCAGTACTTTTAT 51
Db      |||||

RESULT 15
US-10-027-632-284167/c
; Sequence 284167, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284167
; LENGTH: 750
```


QY 1196 AATCCACCGCGCTCGCCCTGCCAAAGTGTGGGATACAGGCGTGAGCCATTGTGCTGG 1255
 Db 4706 TGATCGCCCGCTTAATCTCCCAAGTGTGGGATACAGGCGTGAGCCACGTCGCCGG 4647
 QY 1256 CGAATTTTAAAAAATGATTTCTTAATGTCAGTTTTCATAAGTTTATTTAAATGCAIT 1315
 Db 4646 CC-----ATACAAGTGTCTTTTCAATAATAACTTATTTCCAGGAATTTATCCAT 4595
 QY 1316 TTCCATTGATGTAAGCTTTCAAATTTATAGTATGTGTCTAGTATTTTCTATCTT 1375
 Db 4594 CTCTCTTAGATTTCTAGTCTGTGTGCATGAAGGTGTTCATAGTAG-CTTGAATATCTT 4536
 QY 1376 TTGTAATCTGTTTCAGGCTCTGTAGATGCTCTTTTAAATAATAATATTTTGTGTTG 1435
 Db 4535 TCATATTTCTGTGGTATGTTGTTGTAATATTTCCATTTCTAATGAGCTTATTTG 4476
 QY 1436 CGCTTTTGTATTTTCTTATGCTCTGTGAGAGGATATGCAAAATTTACTAGTGA 1495
 Db 4475 GATCGTCTCCCTTTTCTTGTGTTAATCTCACTAAGGCTCTGCAATTTTGTATCTC 4416
 QY 1496 TCCAAAGATAAATCTTGGGTTGGCAATCTTTTCTCATCTATCTTGTCTTTATATTTA 1555
 Db 4415 TTCAATAACCGATTTTGTTCATTAATCTTTTGTATTTTGTTCATTTAT- 4360
 QY 1556 TTAATCTGTTCTGTTTATAATGCTCTTTTATCTTCTTGTGTTTACTTTGCTGTT 1615
 Db 4359 TTAGTCTGCTCGATCTTGTATTTCTTTCTCTGGGTTTGAGTTTGGTTGTTCTT 4300
 QY 1616 CTTTGTAAATCCTCAGTAGAATGCTTAATTAATCATTTAGCATTCAGTCTTCTTCAATCTA 1675
 Db 4299 GTTCTCTGAGT-----TCCTTGAGGTGTCACCTTAGATGTTGTTCTTCAATGCTGT 4251
 QY 1676 CTATGAGTATTTAGAGCATAAATTTCCCTTTAACTTCCCTTCCACTTCAACTATCATC 1735
 Db 4250 TCCAGACTATTTAATGCTATGAATGTCCTCTTAGTTTGTCTGT-----ATC 4204
 QY 1736 TCACAAATTTGATAGGAGTAGTTTAATATCATTTAGTATCTAAATTAATTTTAAATTTTC 1795
 Db 4203 TCAGAGGTTTAAATAGTGTGTGTCACATTAATTTGTTTCAAGCAATTTTAAATTTTC 4144
 QY 1796 TGATTTTCTTCTGATCTCGCA-----ACTATTTACAAGTATTTTAAATCTCGA 1849
 Db 4143 CTTCTTGAATTCATTTGTTAAACCAAGGTCAITTCAGGAGCAGATTAATTAATTTTCAATG 4084
 QY 1850 ATATAAGATGTTATTTGTTATTTGTTGATCTGATCTCTAAATGAATATATTTAGATC 1909
 Db 4083 ATTTGATAGTTTAAAGGTTCTTTTGGAGTTAAATTTCCAGTTTATTCATTTGCTGTC 4024
 QY 1910 AGATAATGTTGTTTGTAGGACACTAATCTTTGACAAATTTGAGGCTTCTTTTGAAC 1969
 Db 4023 TGAAGAGGTACTTGACATAATTTCAATTTCTTAAATTTGTTGAGACTTGTGTTGGCC 3964
 QY 1970 TAATATGCTCAATTTTATAGACGTTCTGTGTTCTTTGGGAAACATGATTTGAT 2029
 Db 3963 TATCTTATGATTTATCTTGGAGAACGTTCTATGCTGATTTAAATATGATATATCTGA 3904
 QY 2030 GGTGTTTGTGTTAAATTTTGTATTTGTATACATTAGTTTGCAGTTTGTCTTATTTGGCT 2089
 Db 3903 AGTTTGTGGTAGAATGTTAGTAAATATCTGCTAAATCTATTTGTTCTAGGATATTAAT 3844
 QY 2090 GAAATCTCCATTAATCCCTTAATGCTCTCTCATTTGCTGCTTCCCTTTATTAATTAGAG 2149
 Db 3843 AAGTCAATTTTCTTATTTGACTTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3784
 QY 2150 ATAAATGTTAAATTAATCTCACTTATAGTATGATGCTGTTTATATATATATATAA 2209
 Db 3783 AGTATTG--AAGTCCCTACTATTAATTTGTTGGTCCATCTATCCCATTTCTTAGTCTAGT 3726
 QY 2210 ATTTATTAATTCATAAATTTATGTTATGTTAATTTGGAGACCTTAATTAATCATATATAAC 2269
 Db 3725 AATAATTTGTTTAAATTTTGGAGCTCCAGTATTTAGTGCATATATATTTAGATTTGTG 3666

QY 2270 AGAAT-TGTTGATGAATGACAGACTTATATCTATGATAGTACCTTTTATCTCTGTCAT 2328
 Db 3665 ATATTTCTTGTGGACCGATCTTTTATCATTAATAAATTTCTTCTTTTGTCTTTTATA 3606
 QY 2329 AATGTTATT 2337
 Db 3605 ACTGTTGT 3597
 RESULT 2
 US-08-742-185-101
 ; Sequence 101, Application US/08742185
 ; Patent No. 6020476
 ; GENERAL INFORMATION:
 ; APPLICANT: Page, David C.
 ; APPLICANT: Reijo, Renee
 ; APPLICANT: Saxena, Richa
 ; APPLICANT: Hawkins, Trevor
 ; APPLICANT: Reeve, Mary Pat
 ; TITLE OF INVENTION: DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA
 ; NUMBER OF SEQUENCES: 102
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 ; STREET: Two Militia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: US
 ; ZIP: 02173
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/742,185
 ; FILING DATE: 30-OCT-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/690,734
 ; FILING DATE: 31-JUL-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/310,429
 ; FILING DATE: 22-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Granahan, Patricia
 ; REGISTRATION NUMBER: 32,227
 ; REFERENCE/DOCKET NUMBER: WHI94-07A2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 861-6240
 ; TELEFAX: (617) 861-9540
 ; INFORMATION FOR SEQ ID NO: 101:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 43795 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-742-185-101
 Query Match 9.4%; Score 238.2; DB 3; Length 43795;
 Best Local Similarity 50.4%; Pred. No. 4.3e-37;
 Matches 775; Conservative 0; Mismatches 743; Indels 21; Gaps 7;
 QY 912 AATTTTATTTGTTGATATATGACTTTTATTTGTTATATGCTATACAACTAGATTTTTT 971
 Db 16692 ACTCATTAATGTTCTGTTCAGGTTTCTAATTTCTTTTTTTTTTTTTTTTTTTTT 16751
 QY 972 TTTGTTTACCAGATCTTACTCTGTCCACCGCTGAGTGTAAATGGCATGCTCTCAGC 1031
 Db 16752 TTTTGTAGAGGAGTCTCTCTGTCCACCGCTGAGTGTGAGTGTGCAATCTCGGC 16811
 QY 1032 TCACGTCAACTCCGCTCCCGGTTCAAGCCATTCTTCCACCTCAGCCTCCCAAGTAGC 1091

RESULT 4

US-10-096-960-3
; Sequence 3, Application US/10096960
; Patent No. 6664085
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001158DIV
; CURRENT APPLICATION NUMBER: US/10/096,960
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 09/800,960
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 62804
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(62804)
; OTHER INFORMATION: n = A, T, C or G
US-10-096-960-3

Query Match 9.3%; Score 236.4; DB 4; Length 62804;
Best Local Similarity 81.7%; Pred. No. 1e-36;
Matches 273; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 928 TATATGACTTTATTTATATGCTATACAACTAGATTTTTTTTGTGTTTTCACCGAGT 987
Db 43514 TATAATGATATCTTAAAGATTGCTTTCTTTTTTTTTTTTTTTTTTTTTTTTAGACGGA 43573
QY 988 CTCTACTCTGTCACCCAGGCTGGAGTGTATGGCATGTCTCAGCTACTCAACCTCCG 1047
Db 43574 GTCTACTCTCACCAGGCTGGAGTGTATGGCATGTCTTGTGCTACTCAACCTCCG 43633
QY 1048 CTCGGGGTTCAGGCATCTTCCACTCAGCTCCCAAGTAGCTGGGACTACAGGCATG 1107
Db 43634 CTCGGGGTTCAGGCATCTTCCAACTCAGCTCCCAAGTAGCTGGGACTACAGGCATG 43693
QY 1108 AGCCACCGCACCCGGCTAAATTTTGTATTTTATAGAGAGCTGGTTCCTACTATGTGGC 1167
Db 43694 TGCCACACACCCAGCTAAATTTTATATTTTATAGAGAGCTGGTTCCTACTATGTGGC 43753
QY 1168 CAGGCTGATCTGAACTCTGACCTTGTATCCACCGCTCGCCCTGCCAAGTGTGG 1227
Db 43754 CAGGCTGGTCTGATCTCTGACCTTGTATCCACCGCTCGCCCTGCCAAGTGTGG 43813
QY 1228 GATTACAGGCTGAGCCATGTGCTGGCCGATT 1261
Db 43814 GATTACAGGCTGAGCCATGTGCTGGCCGCAATT 43847

RESULT 5

US-09-268-992-7/c
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER FILING DATE: 1999-01-22
; EARLIER FILING DATE: 60/106,056
; EARLIER FILING DATE: 1998-10-28

; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7
Query Match 9.2%; Score 232.8; DB 4; Length 72604;
Best Local Similarity 68.5%; Pred. No. 5.1e-36;
Matches 318; Conservative 2; Mismatches 144; Indels 0; Gaps 0;
QY 805 TAATTTCTTTATTTGTTGTAATGACTATTTTACGTATATTAACAAATTTTATTTGTTG 864
Db 64262 TTATTTAAATTTTTCGATATTTTAAACGTTTCATAAAATGTTATGAGTGTATTTAGTCA 64203
QY 865 TTAAATGACTTTTATTTGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATTTGCT 924
Db 64202 TAATMTACTTAATCATCTCTCTGTTTTCACATTAAGATGTTTAAACATATATAT 64143
QY 925 TGATATATGACTTTTATTTATATGCTATACAACTAGATTTTTTTTGTGTTTGTACCG 984
Db 64142 ATATATATATATATATATATATATATATTTTTTTTTTTTTTTTTTTTTTTTAAACAG 64083
QY 985 AGTCTTACTCTGTACCCAGGCTGGAGTGTATGGAATGTCAGTCTCAGTCTACTGCAACCTC 1044
Db 64082 AGTCTTCTCTGTGCGCCAGGCTGGAGTGTGAGTGGCGCATCTCTGCTACTGCAAGTTC 64023
QY 1045 CGCTCCCGGGTTCAGGCCATTTCTCCACTCAGCTCCCAAGTAGCTGGGACTACAGGC 1104
Db 64022 TGCTCCCGGGTTCAGGCCATTTCTCTGCTCAGCTCCCGAGTAGCTGGGACTACAGGC 63963
QY 1105 ATGAGCCACCGCACCCGGCTAAATTTTGTATTTTATAGTAGAGAGCTGGTTCCTACTATGTT 1164
Db 63962 GCCCGCACCATGCCCGCTAAATTTTATTTTATAGTAGAGAGCTGGTTCCTCGTGT 63903
QY 1165 GGCAGGCTGATCTCGAATCTCTGACTTGTATCCACCGCTCGCCCTGCCAAGTGC 1224
Db 63902 ATCCAGGATGCTCTCAATCTCTGACTCTGATCCGCGCTCAGCTCCCAAGTGC 63843
QY 1225 TGGGATTACAGGCTGAGCCATGTGCTGGCCGATTTTTTTAAA 1268
Db 63842 TGGGATTACAGGCTGAGCCATGTGCTGGCCGCAACTTTTTAATA 63799
RESULT 6
US-09-657-474-7/c
; Sequence 7, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05

```

; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n_positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

```

Query Match	9.2%	Score 232.8;	DB 4;	Length 72604;
Best Local Similarity	68.5%;	Pred. No. 5.1e-36;		
Matches 318;	Conservative 2;	Mismatches 144;	Indels 0;	Gaps 0;
QY	805	TAATTTCTTTATTTGGTTGTAATATGACTATTTTACGTCAATAACAATTTTATTTGTTG 864		
Db	64262	TTAATTTAAATTTTGTGAATATTTTAAAGCTTTTATAAAATGATTAGAGTATTTAGTCA 64203		
QY	865	TTAAATGACTTTATTTGTTGTCATATGATAAATTTTATGTCATAGAACAAATTTTATTTGCT 924		
Db	64202	TAATMTACTTAATCATCTCTCTGTTTTCACATTAAAGAGTAGTTTAAACATATATATAT 64143		
QY	925	TGATATATGACTTTATTTGTTTATATGCTATACAACATAGATTTTTTTTGTGTTTTTGACCG 984		
Db	64142	ATATAATATATATATATATATATATATATATTTTTTTTTTTTTTTTTTTTTTTTAAACAG 64083		
QY	985	AGTCTTACTCTGTCAACCGGCTGGAGTGTAATGGCATGGTCTCAGCTCAGTGCACACCTC 1044		
Db	64082	AGTCTTGCTCTGTGGCCAGGCTGGAGTGACGTGGCGGATCTCTGCTCAGTGCAGATYC 64023		
QY	1045	CGCTCCCGGGTTCAGCCATTCCTCCACCTCAGCCTCCCAAGTAGCTGGGACTACAGGC 1104		
Db	64022	TGCCTCCCGGGTTCAGCCCATTCCTCGCTCAGCCTCCGAGTAGCTGGGACTACAGGC 63963		
QY	1105	ATGAGCACCGCACCGGGCTAATTTTGTATTTTATTTAGTAGAGACGTGGTTCACATATGTT 1164		
Db	63962	GGCCGCCACCATGCCCGGCTAATTTTTTATTTTAGTAGACGCAGATTTCTCGTGTTT 63903		
QY	1165	GGCCAGGCTGATCTCGAACTCTCTGACCTTGTAATCCACCGGCTCGGCTGCACAAAGTC 1224		
Db	63902	ATCCAGGATGGTCTCAATCTCTGACCTCGTATCGGCCGCTCAGCCTCCCAAGTGC 63843		
QY	1225	TGGGATTACAGCGGTGAGCCATTGCGTCGGCCGATTTTTTAAA 1268		
Db	63842	TGGGATTACAGGCATTGAGCACTGGCCGGCCAACTTTTAATA 63799		

RESULT 7
 US-09-685-853A-3
 ; Sequence 3, Application US/09685853A
 ; Patent No. 6479270
 ; GENERAL INFORMATION:
 ; APPLICANT: WEI, Ming-Hui et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
 ; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: CL000871
 ; CURRENT APPLICATION NUMBER: US/09/685,853A
 ; CURRENT FILING DATE: 2002-05-06
 ; PRIOR APPLICATION NUMBER: 60/182,194
 ; PRIOR FILING DATE: 2000-02-14
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 74962
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc feature

```

; LOCATION: (1)....(74962)
; OTHER INFORMATION: n = A,T,C or G
US-09-685-853A-3

Query Match          9.2%   Score 232.6;   DB 4;   Length 74962;
Best Local Similarity 66.9%; Pred. No. 5.6e-36;
Matches 331; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 769 TTGACTCTGTTTGCGAAAAATTTTAAACACTACCTGTTTAAATTTCTTTATTGGTTGTTAAAT 828
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 829 GACTATTTTACGTCATATACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCAT 888
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 889 ATGATAATTTTATGTCATAGAACAAATTTTATTTGTTGATATATGACTTTATTTGTTATAT 948
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 949 GGCTATACAACTAGATTTTTTTTGTGTTTGTACCCAGAGCTTACTCTGTCAACCGAGCTG 1008
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1009 GAGTGAATGGCATGTCTCAGCTCACTCAACCTCCGGCTCCCGGTTTCAAGCCATTCT 1068
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1069 TCCACTCAGCTCCCAAAGTAGCTGGGACTACAGGCATGAGCCACCGCACCCGGCTAATT 1128
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1129 TTGTATTTTAGTAGAGAGTGGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCCTG 1188
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1249 TGCTGGCCGCAATTTT 1263
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

US-09-078-294-9
; Sequence 9, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: BAC-F2 contig 5
US-09-078-294-q

```

	Query Match	9.0%	Score 228.8;	DB 3;	Length 1701;
	Best Local Similarity	78.1%	Pred. No. 1.5e-35;		
	Matches 275;	Conservative	0;	Mismatches 77;	Indels 0;
Qy	914	TTTTTATGCTTGATATATGACCTTATATGTTATATGCTATACAACTAGATTTTTTGT	973		
Db	1057	TCCTCATTTTGATTCATTACTCTCTCTTTATTTTATTTTATTTATTTATTTT	1116		


```
Db 16559 AAAACTTTT 16566

RESULT 11
US-10-193-295-3
; Sequence 3, Application US/10193295
; Patent No. 6620408
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: CL001195DIV
; CURRENT APPLICATION NUMBER: US/10/193,295
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 08/819,993
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 28001
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(28001)
; OTHER INFORMATION: n = A,T,C or G
US-10-193-295-3

Query Match
Best Local Similarity 9.0%; Score 227.6; DB 4; Length 28001;
Matches 287; Conservative 0; Mismatches 79; Indels 2; Gaps 1;

QY 909 AACAAATTTTATGCTTGATATATGACTTATTTATGCTATATGCTATACAACTAGATTTT 968
Db 16199 ATCTATGATGATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16258
QY 969 TTGTTGTTTTGACCGAGTCTTACTGTCTACCCAGGCTGGAGTGTAATGGCATGGTCTC 1028
Db 16259 TTTTGTGAGACGGAGTCTGTCTGTCTACCCAGGCTGGAGTGCACTAATGATGATG 16318
QY 1029 AGCTCACTGCAACTCCGCTCCCGGGTTCAGGCATCTTCCACCTCAGCTCCCAAGT 1088
Db 16319 TGCTCACTGCAAGTCTCTCTCCCGGGTTCAGGCATCTCTCTGCTCAGCTCTGAGT 16378
QY 1089 AGCTGGGACTACAGGCATGAGCCACCGCACCCGGCTAA--TTTGTGATTTTATGATAG 1146
Db 16379 AGCTGGGCTACAGGCATCTGCCACCGACCCGGCTAAATTTTGTATTTTGTAGAG 16438
QY 1147 ACGTGGTTCCACTATGTTGGCCAGGCTGATCTCGAATCTCTGACCTTGTATCCACCCGC 1206
Db 16439 ACGGGGTTTCAACATGTTAGCCAGGATGTTCTCGATCTCTGACCTTGTGATCCGCTGC 16498
QY 1207 CTCGGCTCCCAAGTCTGGGATACAGGCTGAGCAATGTGCTGGCGGATTTTAA 1266
Db 16499 CTCAGCTCCCAAGTCTGAGATTACAGGCATGAGCCACCGCCGCTAATTTATT 16558
QY 1267 AAAAATGT 1274
Db 16559 AAAACTTTT 16566

RESULT 12
US-09-078-294-4
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294

; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contig
US-09-078-294-4

Query Match
Best Local Similarity 8.9%; Score 227.4; DB 3; Length 80246;
Matches 285; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 894 AATTTTATGTCATAGAACAAATTTTATTTGCTTGATATATGACTTTTATTTGTTATATGGCTA 953
Db 38437 AATGGCATGTAATCAAACTGCTACTTATCGTCTAATTCATATCTTGAATGTTTTTATTT 38496
QY 954 TACAACTAGATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1013
Db 38497 TATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 38556
QY 1014 TAATGGCATGCTCTCAGCTCACTGCAACTCGGCTCCGGGTTCAAGCCATTTCTCCAC 1073
Db 38557 CAGTGGGCTGATCTCAGCTCACTGCAACTCGGCTCCGGGTTCAAGCCATTTCTGCTGC 38616
QY 1074 CTCAGCTCCCAAGTAGCTGGGACTTACAGGCATGAGCCACCGCACCGGGTAAATTTTGT 1133
Db 38617 CTCAGCTCCCTGAGTAGCTGGGATTCAGAGGCTGCACTGCACTGCACTGCACTGCACTG 38676
QY 1134 ATTTTATGATAGAGCGTGGTTCCACTATGTTGGCCAGGCTGATCTCGAATCTCTGACCTT 1193
Db 38677 ATTTTATGATAGAGTGGGTTTTCACCATTTGGCCAGGCTGTTGAACTCTTGAATCTTGA 38736
QY 1194 GTAATCCACCGCTCCGCTCCGCAAGTCTGGGATTCAGGCGTGAGCCATTTGTCCT 1253
Db 38737 ATGATCCACCGCTCCGCTCCGCAAGTCTGGGATTCAGGCGTGAGCCATTTGTCCT 1274
QY 1254 GGCCGATTTTAAATAATGT 1274
Db 38797 GGTGCAATGTTTTTATTTATTT 38817

RESULT 13
US-09-146-053-3
; Sequence 3, Application US/09146053A
; Patent No. 6399349
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; APPLICANT: Sprinkle, Terry Joe Curtis
; APPLICANT: Venema, Richard C.
; TITLE OF INVENTION: Human Aminopeptidase P Gene
; FILE REFERENCE: MGI103
; CURRENT APPLICATION NUMBER: US/09/146,053A
; CURRENT FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/057,854
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-053-3

Query Match
Best Local Similarity 8.9%; Score 227.2; DB 4; Length 50000;
Matches 253; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 964 TTTTGTGTTGTTTGTGACCGAGTCTTCTCTCACCAGGCTGGAGTGTATGGCATG 1023
Db 8522 TTTTGTGTTGTTTGTGACCGAGTCTTCTCTCACCAGGCTGGAGTGTATGGCATG 8581
QY 1024 GTCTCAGCTCACTGCAACTCCGCTCCCGGGTTCAAGCCATTTCTTCCACCTCAGCCTCC 1063
```


Db 8582 ATCTCAGCTCACTGCAACCTCCACCTCCCGGTTCAAGCGATTCCCGCTGCCTCAGCGCTCC 8641
Qy 1084 CAAGTAGCTGGGACTACAGGATGAGCCACCGGCTAAATTTTCTATTTAGTA 1143
Db 8642 CGAGTAGCTGGGATTAACAGCGCTGGCCACCGCTTGGCTAAATTTTGTATTTTAA 8701
Qy 1144 GAGACGTGGTTCCACTATGTTGCCAGGCTGATCTCGAACTCTCGACCTTGTAAATCCACC 1203
Db 8702 GAGATGGGTTTACCATGTTGTCAGGCTGCTCGAACTCTCGACCTTGTATCTGCC 8761
Qy 1204 CGCCTCGGCTGCCAAAGTGTGGGATTAACAGCGGTGAGCCATGTCCTGGCCGA 1259
Db 8762 CGCTTGGCTGCCAAAGTGTGGGATTAACAGCGGTGAGCCATGTCCTGGCCGA 8817

RESULT 14

US-09-780-172-18/c
; Sequence 18, Application US/09780172
; Patent No. 6607916
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-ALPHA EXPRESSION
; FILE REFERENCE: RTS-0159
; CURRENT APPLICATION NUMBER: US/09/780,172
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 18
; LENGTH: 63000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-780-172-18

Query Match 8.9%; Score 226.6; DB 4; Length 63000;
Best Local Similarity 81.6%; Pred. No. 7.8e-35;
Matches 262; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
Qy 961 AGATTTTGTGTTTTCACCGAGTCTTACTCTGTCAACCGGCTGAGTGTAAATGGC 1020
Db 43854 AGCCTTTTGTGTTTTCACCGAGTCTTACTCTGTCAACCGGCTGAGTGTAAATGGC 43795
Qy 1021 ATGCTCTCAGCTCACTGCAACCTCCGCTCCCGGTTCAAGCCATCTTCCACCTCAGCC 1080
Db 43794 ATGATCTCAGCTCACTGCAACCTCCGCTCCCGGTTCAAGCCATCTTCCGCTCAGCC 43735
Qy 1081 TCCCAAGTAGCTGGGACTACAGGATGAGCCACCGCAACCGGCTAAATTTTGTATTTTA 1140
Db 43734 TCCCGAGTAGCTGGGATTAACAGGCGCCAGCCACCATGCCAGCTAAATTTTGTATTTTA 43675
Qy 1141 GTAGAGACGTGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCTCGACCTGTATATCC 1200
Db 43674 GTAGAGACAGGGTTTCATCATGTTGGTCAGGATCGTCTCGAACTCTCGACCTCATATCC 43615
Qy 1201 ACCGCTCGGCTGCCAAAGTGTGGGATTAACAGGCGTGAAGCCATGTCCTGGCCGAT 1260
Db 43614 GCCTGCCCTTGGCTTCCAAAGTGTGGGATTAACAGGCGTGAAGCCATGTCCTGGCCGAT 43555
Qy 1261 TTTTAAATAAATGATTTCTTA 1281
Db 43554 TATTATAAGCCCTTCTCTGA 43534

RESULT 15

US-09-816-093-3/c
; Sequence 3, Application US/09816093
; Patent No. 6518055
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001182
; CURRENT APPLICATION NUMBER: US/09/816,093
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 46718
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(46718)
; OTHER INFORMATION: n = A,T,C or G
US-09-816-093-3

Query Match 8.9%; Score 226.2; DB 4; Length 46718;
Best Local Similarity 67.2%; Pred. No. 8.9e-35;
Matches 354; Conservative 0; Mismatches 163; Indels 10; Gaps 2;
Qy 960 TAGATTTTGTGTTTGTGTTTGTGACCGAGTCTTACTGTCTCACCCAGGCTGGAGTGAATGG 1019
Db 15830 TTGTCTTTTGTGTTTGTGTTTGTGACCGAGTCTCGCTCTGT-TGCCAGGCTGGAGTTCAGTGG 15772
Qy 1020 CATGCTCTCAGCTCACTGCAACCTCCGCTCCCGGTTCAAGCCATCTTCCACCTCAGC 1079
Db 15771 CGCGGTCTCGGCTCACTGCAACCTCCGCTCCCGGTTCAAGTGTCTCTCTCCCTCAGC 15712
Qy 1080 CTCCTCAAGTAGCTGGGACTACAGGCAATGAGCCACCGCCGCTAAATTTTCTATTTT 1139
Db 15711 CTCCTCAAGTAGCTGGGATTAACAGGCAACCGCCATGCTGGCTAAATTTTCTATTTT 15652
Qy 1140 AGTAGAGCTGTTTCCACTATGTTGGCCAGGCTGATCTCGAACTCTGACCTGTGAATC 1199
Db 15651 AGTAGAGATGGGTTTCACTATGTTGCCAGGCTGTCTTGTGAACCTTCTGACCTTGTGATC 15592
Qy 1200 CACCCGCTCGGCTCGCAAGTGTGGGATTAACAGGCTGAGCCATGTCCTGGCCG- 1258
Db 15591 CACCCCTCGGCTTACCAAGTGTGGGATTAACAGGCTGAGCCATGTCCTGGCCGCT 15532
Qy 1259 -----ATTTTTAAAAAATGTTTCTTATGTCAGTTTTCATAAGTTTATTTAAAT 1310
Db 15531 ATTTAGGTATTTTAAAGCTGTCTTTCTGTTACTGGATTTTATCTTTTATTTCTGAC 15472
Qy 1311 GCATTTTCCATTTGATGTAAGCTTTCAAATTTATAGTATAGTGTCTCTAGTATTTCTT 1370
Db 15471 ATTGAGAGATCTAACCTCAGAGTTACACATTAGGAAAATTTTATTAAAGATAGTTTCTAT 15412
Qy 1371 ATCTTTTGTAACTGTTTCAGGCTCTGTAGATGTCCTCTTTTAAATAAATAATATTTT 1430
Db 15411 ACCCTAGGTAACACACGCTCAGTTTATACATCTTCCAAACAATAGCTAAAAACGATCA 15352
Qy 1431 GTTTGGCTTTTCTCTATTTTCTTTTCTTATGCTCTTGGAGGGATAT 1477
Db 15351 GCTGCTCTTCTGCTCTGCAATACATTTCTTTTAGAGGGTAT 15305

Search completed: May 21, 2004, 23:07:13
Job time : 202.141 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 16:24:44 ; Search time 1052.27 Seconds
(without alignments)
10258.453 Million cell updates/sec

Title: US-09-914-651A-2_COPY_14040_16580

Perfect score: 2541

Sequence: 1 agcgaattttttttaaatt.....cttctctagtgtgaatttt 2541

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_29Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002s.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	335	13.2	335	3	Aaa94789 Human Pro
2	276	10.9	59215	7	Aad36834 Human tra
3	271.2	10.7	174566	7	Abq77400 Human ITG
4	255.8	10.1	393	4	Aa181268 Human pol
5	252	9.9	157875	6	ABK99972 Human CAD
6	244.2	9.6	9573	9	ADb31452 Bicalutam
7	241.4	9.5	110000	9	ADb11169 2
8	239.8	9.4	21470	8	ABK42270 Genomic B
9	239.8	9.4	21470	8	ADb60426 Connectiv
10	239.4	9.4	16086	4	Aak69721 Human imm
11	239.4	9.4	16086	4	Aal36020 Human mus
12	239.4	9.4	16086	7	ABX59008 cDNA enco
13	238.8	9.4	538	4	Aal04865 Human rep
14	238.8	9.4	538	4	Aal04867 Human rep
15	238.8	9.4	538	4	Aal97761 Human tes
16	238.8	9.4	538	4	ABL97759 Human tes
17	238.4	9.4	10587	3	Aak82382 Human imm
18	238.2	9.4	43795	3	Aaz92583 Human DAZ
19	238.2	9.4	83391	6	ABq67094 Human ang
20	238	9.4	12729	5	AAF97873 Human neu
21	238	9.4	15266	4	AAK73549 Human imm
22	238	9.4	15266	4	Aal37432 Human mus
23	238	9.4	15266	7	ABx60420 cDNA enco

C 24	238	9.4	15271	4	AAK73550 Human imm
C 25	238	9.4	15271	4	AAL37433 Human mus
C 26	238	9.4	15271	7	ABX60421 cDNA enco
C 27	238	9.4	143068	3	AAA34983 Human ade
C 28	238	9.4	143068	3	AAA35150 Human low
C 29	238	9.4	143068	3	AAF21272 Human low
C 30	238	9.4	143068	3	AAF21105 Human low
C 31	238	9.4	143068	6	ABL68124 Ovary can
C 32	238	9.4	143068	6	ABT11034 Human bre
C 33	238	9.4	143068	7	ABZ96966 Human nuc
C 34	238	9.4	143068	7	ABZ96799 Human nuc
C 35	238	9.4	149412	3	AAA35151 Human ade
C 36	238	9.4	152740	3	AAF21273 Human low
C 37	238	9.4	152740	7	ABZ96967 Human nuc
C 38	237.8	9.4	126512	6	ABN83429 Human tra
C 39	237.2	9.3	31952	4	AAK89370 Human dig
C 40	236.4	9.3	62804	6	AAD39317 Human cal
C 41	236.4	9.3	62804	7	ABX10916 Genomic D
C 42	235.6	9.3	17590	3	Aaz50904 Human TBC
C 43	235.4	9.3	185035	6	ABT10147 Human bre
C 44	235.4	9.3	185035	7	ACA64951 Human PEN
C 45	234.8	9.2	13444	4	AAL37295 Human mus

ALIGNMENTS

RESULT 1

AAA94789
ID AAA94789 standard; DNA; 335 BP.

XX
AC AAA94789;

XX
DT 19-JAN-2001 (first entry)

XX
DE Human Prostate-Specific Membrane antigen Enhancer PSME core region.

XX
KW Core region; human; prostate-specific membrane antigen; enhancer; PSME;
KW cancer; renal cell carcinoma; transitional cell; colonic; neuroendocrine;
KW malignant melanoma; pancreatic duct; breast; soft tissue; PSMA;
KW non-small cell lung; testicular embryonal; glioblastoma multiforme;
KW prostate; breast; bladder; cancer; ss.

XX
OS Homo sapiens.

XX
PN WO200052156-A1.

XX
PD 08-SEP-2000.

XX
PF 01-MAR-2000; 2000WO-AU000143.

XX
PR 01-MAR-1999; 99AU-00008956.

XX
PR 25-JAN-2000; 2000AU-00005268.

XX
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX
PI Molloy PL, Watt F;

XX
DR WPI; 2000-594182/56.

XX
PT New regulatory constructs comprising intron 3 of the prostate-specific
PT membrane antigen gene and a heterologous peptide, for directing gene
PT expression in a prostate, bladder, breast or vascular endothelial cell.

XX
PS Claim 10; Fig 11; 56pp; English.

XX
CC The present sequence is the core region from human prostate-specific
CC membrane antigen enhancer (PSME). This sequence is located in intron 3 of
CC the PSMA gene. This sequence was used to construct a recombinant
CC expression cassette, which is useful for the expression of proteins in
CC vascular endothelial cells. In addition, the recombinant construct is
CC useful in the treatment of cancer e.g. renal cell carcinoma, transitional
CC cell carcinoma, colonic adenocarcinoma, neuroendocrine carcinoma,

CC malignant melanoma, pancreatic duct carcinoma, breast carcinoma, soft
 CC tissue carcinoma, non-small cell lung carcinoma, testicular embryonal
 CC carcinoma, glioblastoma multiforme, prostate, breast or bladder cancer
 XX
 SQ Sequence 335 BP; 84 A; 55 C; 54 G; 142 T; 0 U; 0 Other;
 Query Match 13.2%; Score 335; DB 3; Length 335;
 Best Local Similarity 100.0%; Pred. No. 4.6e-40;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 721 AATTATTTTCTCTTAACCTTCAAACTCAAGAAACAGTTGGCTTGACTCTGTTT 780
 DB 1 AATTATTTTCTCTTAACCTTCAAACTCAAGAAACAGTTGGCTTGACTCTGTTT 60
 QY 781 GTGGAATAATTTAAACTACTGTTTAAATTTCTTTATTGTTGTAATATGACTATTTTACG 840
 DB 61 GTGGAATAATTTAAACTACTGTTTAAATTTCTTTATTGTTGTAATATGACTATTTTACG 120
 QY 841 TCATATAACAATTTTATTTGTTTAAATGACTTATTTGTTGTCATATGATAAATTTTA 900
 DB 121 TCATATAACAATTTTATTTGTTTAAATGACTTATTTGTTGTCATATGATAAATTTTA 180
 QY 901 TGTTCATAGAACATTTTATTTGTTTAAATGACTTATTTGTTGTCATATGATAAATTTTA 960
 DB 181 TGTTCATAGAACATTTTATTTGTTTAAATGACTTATTTGTTGTCATATGATAAATTTTA 240
 QY 961 AGATTTTGTGTTTGTGTTTGAACGAGTCTTACTCTGTACCCAGGCTGGAGTGAATGGC 1020
 DB 241 AGATTTTGTGTTTGTGTTTGAACGAGTCTTACTCTGTACCCAGGCTGGAGTGAATGGC 300
 QY 1021 ATGGTCTCAGCTCACTGCAACCTCCGCTCCCGG 1055
 DB 301 ATGGTCTCAGCTCACTGCAACCTCCGCTCCCGG 335

RESULT 2
 AAD36834/c
 ID AAD36834 standard; DNA; 59215 BP.

XX AAD36834;
 XX 07-MAR-2003 (first entry)
 XX Human transporter protein genomic DNA.
 XX Human; transporter; sodium-glucose cotransporter; chromosome 16;
 KW drug screening; single nucleotide polymorphism; SNP; ds.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH 3000..57168
 FT CDS
 FT /tag= a
 FT /product= "Transporter protein"
 FT /note= "The CDS is interrupted by 12 introns"
 FT 3000..3014
 FT /tag= b
 FT /label= 1
 FT 3015..10042
 FT /tag= c
 FT /label= 1
 FT variation replace(4429,A)
 FT /tag= aa
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(5167,A)
 FT /tag= ab
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(6198,C)
 FT /tag= ac
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(6760,A)
 FT /tag= ad
 FT /standard_name= "Single nucleotide polymorphism"

FT variation replace(6934, A/-)
 FT /tag= ae
 FT /standard_name= "Single nucleotide polymorphism"
 FT /note= "The nucleotide is replaced by A or no nucleotide"
 FT variation replace(7038,C)
 FT /tag= af
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(7116,T)
 FT /tag= ag
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(7978,G)
 FT /tag= ah
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(9973,A)
 FT /tag= ai
 FT /standard_name= "Single nucleotide polymorphism"
 FT exon 10043..10147
 FT /tag= d
 FT /label= 2
 FT 10148..12305
 FT /tag= e
 FT /label= 2
 FT variation replace(11318,T)
 FT /tag= aj
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(11998,T)
 FT /tag= ak
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(12047,T)
 FT /tag= al
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(12058,T)
 FT /tag= am
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(12141,G)
 FT /tag= an
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(12163,T)
 FT /tag= ao
 FT /standard_name= "Single nucleotide polymorphism"
 FT exon 12306..12365
 FT /tag= f
 FT /label= 3
 FT 12366..21408
 FT /tag= g
 FT /label= 3
 FT variation replace(19275,G)
 FT /tag= ap
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(2110,A)
 FT /tag= ar
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(21118,C)
 FT /tag= as
 FT /standard_name= "Single nucleotide polymorphism"
 FT exon 21409..21513
 FT /tag= h
 FT /label= 4
 FT 21514..23038
 FT /tag= i
 FT /label= 4
 FT variation replace(22575,A)
 FT /tag= at
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(22914,T)
 FT /tag= au
 FT /standard_name= "Single nucleotide polymorphism"
 FT exon 23039..23144
 FT /tag= j
 FT /label= 5
 FT 23145..23106
 FT /tag= av
 FT /standard_name= "Single nucleotide polymorphism"

```
FT /note= "This variation brings about Val to Ala change at  
FT position 118 in the corresponding protein sequence"  
FT 23145. .29697  
FT /tag= k  
FT /label= 5  
FT /tag= aw  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(24115,T)  
FT /tag= ax  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(24181,A)  
FT /tag= ay  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(24784,G)  
FT /tag= az  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(24784,G)  
FT /tag= ba  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(27236,C)  
FT /tag= bb  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(28095,C)  
FT /tag= bc  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(28329,C)  
FT /tag= bd  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(28584,C)  
FT /tag= be  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(28595,G)  
FT /tag= bf  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(28906, -)  
FT /tag= bg  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(28906, -)  
FT /tag= bh  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(30674, G/A)  
FT /tag= bi  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(31244,T)  
FT /tag= bj  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(32643,T)  
FT /tag= bk  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(32817,T)  
FT /tag= bl  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(32908,C)  
FT /tag= bm  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(32983,C)  
FT /tag= bn  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(33001,A)  
FT /tag= bo  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(33292,C)  
FT /tag= bp  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(33470,T)  
FT /tag= bq  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(33724,A)  
FT /tag= br  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(33803,A)  
FT /tag= bs  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(33834,A)  
FT /tag= bt  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(34736, A/G)  
FT /tag= bu  
FT /standard_name= "Single nucleotide polymorphism"  
FT /note= "The nucleotide is replaced by nucleotides A or G"  
FT replace(36513,T)  
FT /tag= bv
```

```
FT variation replace(33559,T)  
FT /tag= bq  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(33724,A)  
FT /tag= br  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(33803,A)  
FT /tag= bs  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(33834,A)  
FT /tag= bt  
FT /standard_name= "Single nucleotide polymorphism"  
FT replace(34736, A/G)  
FT /tag= bu  
FT /standard_name= "Single nucleotide polymorphism"  
FT /note= "The nucleotide is replaced by nucleotides A or G"  
FT replace(36513,T)  
FT /tag= bv
```

Query Match 10.9%; Score 276; DB 7; Length 59215;
Best Local Similarity 55.2%; Pred. No. 1.1e-31;
Matches 797; Conservative 0; Mismatches 590; Indels 57; Gaps 11;

Qy	915	TTTTATTGCTTGATATGACTTTATTGTTATATGGCTATACAACTAGATTTTTTGTG	974
Db	8542	TCATTTTCTTTGTGAGAAAGTTTCTGATTACTAATTCACTGCTTTACTTTTTTTTTT	8483
Qy	975	TTTTTGACCGAGTCTTACTCTGTCACCCAGGCTGGAGTGAATGCGATGCTCAGCTCA	1034
Db	8482	TTTGAATGGAGTCTTCTTTGTCCACCCAGGCTGGAATGCAATGGTGCATCTCAGCTCA	8423
Qy	1035	CTGCAACTCCGCTCCCGGTTCAAGCCATTCTTCCACCTCAGCTCCCAAGTAGCTGG	1094
Db	8422	CTGCAACTCTGCTTCTCGGTTCAAGTATCTCTATCTCAGCTCTCTGAGTAGCCGG	8363
Qy	1095	GACTACAGGATGAGCCACCGCTAAATTTTGTATTTTGTATTTTAGTAGAGAGCTGTT	1154
Db	8362	GATTACAAGCACATGCCACCAACAGCTAAATTTTGTATTTTAGTAGAAATGGGTT	8303
Qy	1155	CCACTATGTTGGCCAGGCTGATCTCGAACTCTCGA--CCTGTGAATCCACCGCTCGGC	1212
Db	8302	TCACCGGTGTTGCCAGGGTGTCTCAAACTCATGAGTTCAGGTGATTCGCCCTCTCAGC	8243
Qy	1213	CTGCCAAGTCTGGGATTACAGCGGTGAGCATTTGCGCTGGCCGATTTT-----	1264
Db	8242	CTCCCAAGTCTAGGATTAACAGGGTGAGCCACTGCTTGCCCTCTCTTACTTTTAA	8183
Qy	1265	-----TAAAAAATGATTTCTTATGTCAGTTTTCATAAGTTTATTTAAAA	1309
Db	8182	TAGATATAATTCAGGTTTCTATTCTTACTTGAAGTCAAGTTTGTGCTCTGTCTAGGA	8123
Qy	1310	TGCATTTTCCATTTGATGAAGCTTTCAAATTTATAGTATAGTTGTTCCCTAGTATTTCT	1369
Db	8122	ATTGTTTCATTTTCATGTAGTTGTTTAAATTTTGTCAATAATTTGCTCAGCAGCATTTCT	8063
Qy	1370	TA-----TCITTTCTAATCTGTCAGCGTCTGATGCTGCTCTTTTATAATAATATAT	1425
Db	8062	TATACTATATTTAAATTCATATGTTGATGAATATCCCTCTTTTCAITTTCTAATTT	8003
Qy	1426	TATTGTTGCGCTTTTGTCTATTTTCTTATTTGCTCTGTAGAGGATATGTCAAATTT	1485
Db	8002	TATTAATTTGCATCTTCTGCTTTGTTTCTTA--ATCTAGGTAGAAATTTGTCATTT	7946
Qy	1486	TACTAGTGTATCAAAGATAAATTTGG---CGTTGGCAATCTTTTCTCATCTATCTTT	1542
Db	7945	TGTTGATGTTTCAAAGAACAACTTTTGAATTTTGTATTTTCTCTAACAGTTTCTAT	7886
Qy	1543	GCTTTATATTTTATTAATCTGTTCTGTTTATATATGCTCTTTAT--CTTCTTTGT	1600
Db	7885	TCATATTTTGTGTTTACCTCCTAATCTTTGTTATTTTCTTCTGCTGCTTTGCTTTGG	7826
Qy	1601	GTTTACTTTGCTGTTCTTTGTAAAAATCCTCAGT-AGAAATGCTTAATTTATTCATTCAG	1659

Db 7825 GTTAAATTTGCCCTCTCTTTTAGTTTCTTAATGTAAAGATGAGGTAATGTTGGACA 7766
QY 1660 TCCTTCTCTACTATAGATATTTAGAGCAATAAATTTCCCTTTAACTTCCTCTT 1719
Db 7765 TCCTTCTCTCTTTTAAATGGCAATGGTGAAGTTTCTTCTAAGCACTGCTTTAG 7706
QY 1720 CCATTTCAACTCACTCAAAATTTGGATTAGGAGTAGTTTAAATATCATATAGTATCTA 1779
Db 7705 CTGGATCCCATAGTTTTCAGCATGTG-----TTTTCATTTTCAGTAATCTCAA 7657
QY 1780 AATATTTTAAATTCGTATTTCTCTTTCCTGCTGCACTATTTACAGATATTTT 1839
Db 7656 AGTATTTTCTAAATTCGTGTGATTTCTTCTTGGCCCAATGGTTATTTAGAGGTGT 7597
QY 1840 AAAATCTCAATATAAGATTTGTTATTTGTTTGAATCTGATCTTAAATTCGAATA 1899
Db 7596 TGTTTAAATGATATGTTTGGAAATTTTCCAAGTTATTT---ATTATAATTTTCATC 7540
QY 1900 TATTTAGATCAGATAATGTTGTTGTAGGACACTAATCTCTTGCACATTTGTTGAGGCTTC 1959
Db 7539 CACTGTGGATAGACAACACTTTGTATTAATTTCAATCATTTTAAATTTCTTGA-GATTT 7481
QY 1960 CTTTGGAACTATATCTGCTCAATTTTATAGAGCTTCTGTTCTTTTGGGAAACA 2019
Db 7480 ATTTATGGCTTAAACATACAGTCTATCTGGAGATGTTCCATGTGCACCTTAAGATGAATA 7421
QY 2020 TGTATTTGATGGTTGTTGGTTTAAATTTGTTATTTGTATTTGTACATTTAGTTTGCTTA 2079
Db 7420 GGTATTTTGTGTTGTTGGATTAATTTTATATATGTTCTGTAGTCACTGGTTTAT 7361
QY 2080 TTATTTGGGTGAATCTCCATTAATCTTAATGCTCTCTCATTTTGTCTGCTTCCCTTA 2139
Db 7360 AGTGTGTTCACTGCTCTTCTATTTTCTGTTGATCTTATTTTGTTTTCCATACATTTATG 7301
QY 2140 TTAATTTAGATATAATGTTAAATTAATCTCACTTATAGTATGCTCTGTTTATACTA 2199
Db 7300 AAAATGAGATTTAAAGTCCCAATTAATAGTGT---TTGAAATGCTATTTTCTCCCT 7245
QY 2200 TATATATAAATTTAATTTCCATAATTTATGTTATGTTATGTTATGTTGAGACCTTATATC 2259
Db 7244 CATTTGTCAGTTTGTCTCATATATTTGGGAGTCTGTTGTAGTGCAATATGTTT 7185
QY 2260 ATATATAACAGAAATTTGATGAATGACAGACTTATATCTATGTTAGTACCTTTTATA 2319
Db 7184 ATAATTTGTTATGCTCTTCTGAGGATGACCTTTTATCATATGTTAAATTTCTTTTTC 7125
QY 2320 TCTC 2323
Db 7124 TCTC 7121

RESULT 3
ID ABQ77400/c standard; DNA; 174566 BP.

XX AC ABQ77400;

XX DT 10-MAY-2003 (first entry)

XX DE Human ITGB3 DNA.

XX Human; ITGB3; vascular disease; cardiatic; antiarteriosclerotic; stroke;
XX cerebroprotective; gene therapy; coronary artery disease; ischaemia;
XX myocardial infarction; peripheral vascular disease; pulmonary embolism;
XX venous thromboembolism; forensic; paternity testing; GI14589612; gene;
XX SNP; single nucleotide polymorphism; ds.

XX OS Homo sapiens.

XX Key Location/Qualifiers
XX variation replace(107078,g)
XX FT /*tag=a
XX FT /standard_name= "SNP"

FT FT /note= "single nucleotide polymorphism (ID ITGB3u2) which
FT results in a change from a Leu to Pro in the ITGB3
XX protein at residue 59"

XX PN WO2003016494-A2.

XX PD 27-FEB-2003.

XX PF 16-AUG-2002; 2002WO-US026343.

XX PR 16-AUG-2001; 2001US-0313097P.

XX PR 05-OCT-2001; 2001US-0327485P.

XX PR 14-DEC-2001; 2001US-00020141.

XX PA (VITI-) VITIVITY INC.

XX PI McCarthy J, Ableson A;

XX DR WPI; 2003-300617/29.

XX DR P-PSDB; ABG74668.

XX PT Identifying a subject as a candidate for a particular course of therapy
XX to treat a vascular disease or disorder, e.g. stroke, myocardial
XX infarction or ischemia by determining the identity of the nucleotide
XX present at specific positions.

XX PS Claim 1; Fig 1; 568pp; English.

XX CC This invention describes a novel method for identifying a subject as a
XX candidate for a particular course of therapy to treat a vascular disease
XX or disorder. The method comprises determining the identity of the
XX nucleotide present at specific positions, or their complements, and
XX identifying the subject as a candidate for a particular clinical course
XX of therapy based on the identity of the nucleotide present in that
XX specific position. The method can be used for identifying a subject who
XX is a candidate for further diagnostic evaluation of a vascular disease or
XX disorder and selecting a clinical course of therapy. The products of the
XX invention have cardiac, antiarteriosclerotic and cerebroprotective
XX activity and can be used for gene therapy. The methods disclosed are
XX useful for treating a vascular disease, e.g. atherosclerosis, coronary
XX artery disease, myocardial infarction, ischaemia, stroke, peripheral
XX vascular diseases, venous thromboembolism and pulmonary embolism. The DNA
XX sequences are useful as fingerprint for detecting different individuals
XX within the same species applicable in forensic studies and paternity
XX testing. This sequence encodes the human ITGB3 gene represented in
XX CC GI14589612, used to illustrate the method of the invention

XX SQ Sequence 174566 BP; 53756 A; 36885 C; 36182 G; 45895 T; 0 U; 48 Other;

Query Match 10.7%; Score 271.2; DB 7; Length 174566;
Best Local Similarity 53.0%; Pred. No. 5e-31;
Matches 802; Conservative 0; Mismatches 678; Indels 34; Gaps 9;

QY 880 GTTGTGTCATGATGATAATTTTATGTCATAGAACAAATTTTATGCTTGATATATGACTTA 939

Db 37228 GTTGTGTCATGATGAGCCCTTTATTTATTTGAGGTAGTTTCTTTTATAGTGAATCAGATTG 37169

QY 940 TTGTTATATGGCTATA-CAACTAGATTTTTTGTGTTTGTGACCGAGTCTTACTCTGTC 998

Db 37168 TATATTTATTTATTTATTTATTTTATTTTTCAGATGGAGTCTTGTGTCGTC 37109

QY 999 ACCAGGCTGGAGTGAATGGCAATGCTCAGCTCACTGCAACTCCGCTCCCGGCTTC 1058

Db 37108 ATCAGGCTGGAGTGCAGTGCACAAATCTTGGCTCACTGCAACTCTGCTCCCGGCTTC 37049

QY 1059 AAGCCATTTCTCCACCTCAGCCTCCCAAGTAGCTGGGACTACAGGATGAGCCACCCGAC 1118

Db 37048 AAGGATTTCTTCCCTCAGCCTCCCAAGTAACTGGTACTACAGGATGACCACTATGC 36989

QY 1119 CCGGCT-AAATTTTGTATTTTATAGTACGAGTGGTTCCTATGTTGGCCAGGCTATC 1177

Db 36988 CTGGCTAAATTTTGTATTTTATAGTACGAGTGGGTTTACCATGTTGGCCAGGCTATG 36929

QY	1178	TGGAACCTCTGACCT--GGTAATCCACCCGCTCGGCTGCCAAAGTGCTGGGATTACAG	1235
DB	36928	TCGAACCTCTGACCTCAAGTGATCGCGCTGCCCTTGCTCCCAAAGTGTGGAATTACAG	36869
QY	1236	GGGTGAGCCATTGCTGCTGCCGATTTTTTAAATAATGTAATCTTATGTCAAGTTTTCATA	1295
DB	36868	GTGTGAGCCACTGCACCCAGCCAGATTTTATTTGTGTAGAATCAAGSATATGTGAATGT	36809
QY	1296	AGTTTATTTAAATGCAATTTTCCATTGATGTAAGCTTTCAAAATTTATAGTATAGTTGT	1355
DB	36808	GATAGTCTTTATAGAACTGGAATAGCAAGTAAGGGCAATAAAAACCTCTTCAATACCTAT	36749
QY	1356	TCCTAGTATTTCTTATCTTTTGTAATCTGTTCAGCGTCTGTAGATGCGCTCTTTTAA	1415
DB	36748	T-----TACTCAATTTATGTAATAAACATTAATATGCCATTATATGACAGTGAACAA	36696
QY	1416	TAAATAATTAATTTGTTTTGCGCTTTGCTATTTTTTTTCTTATTGCTCTTTGAGGGAT	1475
DB	36695	TATAGATGATGCTTGGCTTGCTTTTCTTTTGGCAATTTTGATAGACTCCATGGCAATT	36636
QY	1476	ATGTCAAAATTTACTAGTGTATCCAAAGATAAACTTTGGGTTGGCAATCTTTTCTCATC	1535
DB	36635	AGGA-----TTTTCTCGTCTCTTTTATCTACTTGGTTTTTGACTACATTTTTTTTC	36582
QY	1536	TATCTTTGCTTTATATTTTAAATCTCTGTTCTGTGTTTTTAAATGCTCTTTTATCTTC	1595
DB	36581	TGTTTTCAGTTTTCATGTAGTCTCTCTCAATCTTGGTTATTTCTTTCTTCTGCTGGTT	36522
QY	1596	TTTGTGTTTTACTTTTCTGTTCTTTGTGTGTAATAATCTCTAGTAGAATGCTTAATGACAT	1655
DB	36521	TGGGTTTGGTTTGTCTTGTCTCTAGTACCTTGAGGTTGACCTTAGATTGCCAATTT	36462
QY	1656	TCAGTCTTTCTTCAATTTCTACTATGAGTATTTAGAGCCATAAATTTCCCTTTAACTTCC	1715
DB	36461	GTGCTCTTTTCAGACTTTTGTAGTAGTCAATTTATGCTAATTAATTTCTCTCTTAA	36407
QY	1716	CTTTCCACTTCAACTACATCTCACAAATTTGGATTAGGATGATTTAAATTAATCATTAGTA	1775
DB	36406	--TACTGCTTTTGTCTGTATCCAGAGGTTTGATAGTTTGTGCTACTAATTAATGTTTCACT	36349
QY	1776	TCTAAATATTTTAAATTTCTGTAATTTCTTCTTTGATCCTGCACTATTTTACAATTT	1835
DB	36348	TCAAAGATTTTAAATTTCCATCTGTATTTCAATTTGACCCAGTGATTTTTCAGGAG	36289
QY	1836	TTTTAAATCTCGAATAAAGATTTGTTATCTGTATTTGTTTGTCTGATCTCTAAATTTG	1895
DB	36288	AGATTATTTAAATTTCCCTGTATTTTGCAATGTTTTCAGGGTTCCTTTTGAGTTGATTTCC	36229
QY	1896	AATATA--TTGATACAGATAATGCTGGTTGTAGGACACTAACTCTTTGCAATTTGTGAG	1954
DB	36228	AGTTTATTTCACTGTAGCAAGAGTACTTGCTGTAATTTTAAATTTCTTAAAATTTGTTGAG	36169
QY	1955	GCTTCTTTTGGAACTAATATGCTGCTCAATTTTATAGACGTTCTGTGTTTCTTTGGGAA	2014
DB	36168	GCTTGTTTTGGGCTACCATATGGTCTATCTTGAGGATGTTTCCATGTGCTGATGAATA	36109
QY	2015	AAACATGTAATTTGATGTTGTTTGGTTTAAATTTTGTATTTGTACATAGTTTTCAGTTT	2074
DB	36108	GAATTCATAGTCTCAATTTATTGGGTAGAAATGTTCTGTAAATTTCTGTTAGATCAATTT	36049
QY	2075	GCTTATTTTGGCTGAAATCTCCATTAATCTTAATGCTCTCTCAATTTGTCTGCTTC	2134
DB	36048	TTTTAGAGTATAGATTTTAACTGCAATTTTCTTTGTTG-----ACTTCTGCTTGATG	35996
QY	2135	CTTTATTAATTTAGAGATAAATGTTAAATTAATCTCACCTCACTATAGTCAATGCTGTTTAA	2194
DB	35995	ACCTGTCTAATGCTGTCAATTTGAGTATTGAAGCCCCCTACTACTGTGCTGCTCTCTCA	35936
QY	2195	TACTATATATAAAATTTTAAATTTCCATAAATTTATGTTATGATAAATTTTGAGACCTA	2254
DB	35935	TTTCTTAGGCTAGTAGTAATCTGTTTATAAATTTTGGAGCTCCAGTGTTAGTGTG--CAT	35878
QY	2255	TTATCATATATAAACAATTTGATGAAATGACAGACTTATACCTATGTAGTAGTCTT	2314

Db	35877	ATATGTTTAGGATTACGATATTTCTGTTGGACTAGTCCTATATTATATAATATGTCCT	35818
QY	2315	TTTTTATCTCGTCATAAATGTTATTTGACTTTGTCTCTAAATTTTTTTTTTAATTAATATTGT	2374
Db	35817	CTTTGGCTTTTAAACTGCTATTGCTTTAAAGTTTGTGTTCTGATACAGAAAGCTAC	35758
QY	2375	TTGGTATTTCTTTT	2388
Db	35757	TCCTGCTTGCTTTT	35744
RESULT 4			
AAI81268			
ID	AAI81268	standard; cDNA; 393 BP.	
XX			
AC	AAI81268;		
XX			
DT	06-NOV-2001	(first entry)	
XX			
DE	Human polynucleotide SEQ ID NO 1328.		
XX			
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
KW	nervous system disorders; arthritis; inflammation; ss.		
XX			
OS	Homo sapiens.		
XX			
FN	WO200164835-A2.		
XX			
PD	07-SEP-2001.		
XX			
PF	26-FEB-2001; 2001WO-US004927.		
XX			
PR	28-FEB-2000; 2000US-00515126.		
PR	18-MAY-2000; 2000US-00577409.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
FI	Tang YT, Liu C, Drmanac RT;		
XX			
DR	WPI; 2001-514838/56.		
DR	P-PSDB; AAO01337.		
XX			
PT	Isolated nucleic acids and polypeptides, useful for preventing diagnosing		
PT	and treating e.g. leukemia, inflammation and immune disorders.		
XX			
PS	Claim 1; SEQ ID NO 1328; 1399pp + Sequence Listing; English.		
XX			
CC	The invention relates to human polynucleotides (AAI79941-AAI93841) and		
CC	the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to		
CC	cytokine, cell proliferation or cell differentiation or which may induce		
CC	production of other cytokines in other cell populations. The		
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or		
CC	peptide therapy. The polypeptides have various cytokine-like activities,		
CC	e.g. stem cell growth factor activity, haematopoiesis regulating		
CC	activity, tissue growth factor activity, immunomodulatory activity and		
CC	activin/inhibin activity and may be useful in the diagnosis and/or		
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and		
CC	inflammation. Note: The sequence data for this patent did not form part		
CC	of the printed specification, but was obtained in electronic format		
CC	directly from WIPO at ftp.wipo.int/pub/published_pct_sequences		
XX			
SQ	Sequence 393 BP; 102 A; 56 C; 71 G; 163 T; 0 U; 1 Other;		
Query Match	10.1%;	Score 255.8; DB 4; Length 393;	
Best Local Similarity	99.2%;	Pred. No. 1.4e-28;	
Matches 257; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
QY	1919	GGTTTGTAGGACACTAATCTCTTTGACAAATGTTGAGGCTTCTTTGGAACTAATATG	1978
Db	12	GGTTTGTAGGACACTAATCTCTTTGACAAATGTTGAGGCTTCTTTGGAACTAATATG	

QY 1979 CTCAAATTTTATAGAGCTTCTGTTTCTTTGGAACAAATGATATTGATGTTGTTG 2038
Db 72 CTCAAATTTTATAGAGCTTCTGTTTCTTTGGAACAAATGATATTGATGTTGTTG 131
QY 2039 GTTTAATATTTGATTTGATACATGATGTTGAGTTTGGCTTATTTGCTGAATCTCC 2098
Db 132 GTTTAATATTTGATTTGATACATGATGTTGAGTTTGGCTTATTTGCTGAATCTCC 191
QY 2099 ATATTCCTTAAGTGTCTCTCATTTTGTCTGCTTCTTTATTAATAGAGATAAATGTT 2158
Db 192 ATATTCCTTAAGTGTCTCTCATTTTGTCTGCTTCTTTATTAATAGAGATAAATGTT 251
QY 2159 AAATTATCTCACCTCACTA 2177
Db 252 AAATTATCTCACCTCGCCA 270
RESULT 5
ID ABK99972
XX ABK99972 standard; DNA; 157875 BP.
AC ABK99972;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human CADPKL genomic DNA.
XX
KW Human; calcium/calmodulin-dependent protein kinase-like gene; CADPKL; ds;
KW gene; neuropsychiatric disorder; attention deficit disorder; ADD;
KW schizoaffective disorder; bipolar disorder; unipolar affective disorder;
KW schizophrenia; adolescent conduct disorder; pharmacogenomics;
KW fingerprinting; paternity testing; antidepressant; neuroleptic.
XX
OS Homo sapiens.
XX
PN WO200254939-A2.
XX
PD 18-JUL-2002.
XX
PF 07-JAN-2002; 2002WO-US000367.
XX
PR 09-JAN-2001; 2001US-00757300.
PR 23-AUG-2001; 2001US-00935464.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyer JM, Barrington-Martin R, Parker A;
PI WPI; 2002-590643/63.
XX
XX New variants of calcium/calmodulin-dependent protein kinase-like nucleic
XX acids and polypeptides, useful for diagnosing and treating
PT neuropsychiatric disorders, e.g. schizophrenia, schizoaffective disorder,
PT and bipolar disorder.
XX
XX Claim 1; Page 119-200; 223pp; English.
PS
XX
CC The invention relates to a nucleic acid comprising a polymorphic region
CC of a calcium/calmodulin-dependent protein kinase-like gene (CADPKL)
CC allelic variant, and the polypeptide it encodes. CADPKL allelic variants
CC are useful in determining whether a subject has or is at risk of
CC developing a neuropsychiatric disorder, such as schizophrenia, attention
CC deficit disorder (ADD), schizoaffective disorder, bipolar disorder,
CC unipolar affective disorder and adolescent conduct disorder. The
CC polypeptides, polynucleotides, antibodies and modulators of the CADPKL
CC allelic variants are useful for diagnosing or treating these
CC neuropsychiatric disorders. The polypeptides may be used to raise
CC antibodies to a CADPKL polypeptide. The nucleic acids may be used as
CC probes or primers, in pharmacogenomics for designing therapies for the
CC disorders, and in fingerprinting for detection of different individuals
CC with the same species (e.g. paternity testing). This sequence represents
CC human CADPKL genomic DNA of the invention

XX Sequence 157875 BP; 44403 A; 35734 C; 34322 G; 43416 T; 0 U; 0 Other;
SQ
Query Match 9.9%; Score 252; DB 6; Length 157875;
Best Local Similarity 54.0%; Pred. No. 3.1e-28;
Matches 862; Conservative 0; Mismatches 635; Indels 99; Gaps 13;
QY 805 TAAATTTCTTTATTTGGTTGTAATAGACTATTTTACGTATATAACAATTTTATTTCTTTG 864
Db 33292 TAGGCTTTTCTTTGATGGGAGACATTTTATTAAGATTTTAAATCTCTTTACTATAAATTG 33351
QY 865 TTAATAGACTTTTATTTGTCATATGATAAATTTTATGTATAGAAACAATTTTATTTGCT 924
Db 33352 TCTGTTTCAATTTTCTATTTCTTTGTTTCTATCTTTGTTAGTTCGAATACATTTGTC 33411
QY 925 TGAATATAGACTTTTATTTGTTATATGCTATACAACTAGATTTTATTTGTTTGTGTTGACCG 984
Db 33412 CACGAATTTATCAACTTCTGCTAGGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 33471
QY 985 AGTCTTACTCTGTCACCCAG-----GCTGGAGTGTAAATGGCATGCTCTCAGCTCACTGC 1038
Db 33472 AGTCTGACTTTGTTGCCAGGCTGGAGTGGAGTGCAGTCCCGTATCTCAGCTCACTGC 33531
QY 1039 AACCTCCGCTCCCGGTTTCAAGCCATTTCTTCCACCTCAGCCCTCCCAAGTAGCTGGGACT 1098
Db 33532 AAGCTGGCTCTCCAGGTTTCAAGCCATTTCTCTGCTCAGCTCCCGAGTAGCTGGGACT 33591
QY 1099 --ACAGGATAGCCACCGCACCCGGCTAAATTTT-----GTATTTTGTAGAGACGTGG 1152
Db 33592 ACACAGGCGCGCCACCATGCTGCTGCTAAATTTTGTATTTTATTTTGTAGAGATGGGG 33651
QY 1153 TTCCACTATGTTGGCAGCTGATCTCGAATCTCTGACCTTGAATCCACCCGCTCGGC 1212
Db 33652 TTTCACTGTGTTAGCCAGGATGCTCTGATCCCTGACCTCGTGAATCTGCCGCTCGGC 33711
QY 1213 CTGCCAAAGTGTGGGATTACAGCGGTGAGCCATTTGCTGCGCGCATTTTAAATAAAT 1272
Db 33712 CTCCAAAGTGTGGGATTACAGCGGTGAGCCATTTGCTGCGCGCATTTTAAATAAAT 1332
QY 1273 GTATTTCTATGTCAGTTTTCATAAGTTTATTTTAAATGCAATTTTCCATTTGATGTAAGC 1332
Db 33772 TTCTAAATTAATGGCAAT-----TAGTTGTTCAATAGTAGTCTATGATTTCTTTGTTG 33789
QY 1333 TTTCAAAATTTATAGTATAGTTGTTCTAGTATTTTCTTATCTTTGTAATCTGTTCAAGC 1392
Db 33790 -----TAGTTGTTCAATAGTAGTCTATGATTTCTTTGTTGTTGCAATGT 33833
QY 1393 TCTGTAGATGCTGCTCTTTTAAATAAATAATTTATTTGTTGCTGCTTTT--GCTATTTT 1450
Db 33834 TAAATGTAATGTCCTCTTTTATTTTGTAGTTTGTGACTTTTATTTCAATTTGATCTTTT 33893
QY 1451 TTTTCTTTATTTGCTCTTGAGAGGATATGTCAAATTTTACTAGTATGTCAAAGAATAAAT 1510
Db 33894 TTTCTTGTTAAATCTTTCTAATGTTGTCAAATTTTCTTATCTTTTCAAAAAGCAACT 33953
QY 1511 TTGGCGTTGGCAATCTTTTCTCATCTATCTTTGCTTTATTTTATTTAAATCTGTTCTG 1570
Db 33954 TTTTCTTTGATCTTTTAAATAATTTTATAGACTATATTTGTTGATTTCTGCTCCTATCTTT 34013
QY 1571 TTTTATAATTTGCTCTTTTATCTTTTGTGTTTACTTTGTTGTTTCTTTGTTAAATCTC 1630
Db 34014 ATTAATTTCTTCTCTTTTAAATTTTTC--GATTTTGGTTTCTCGTTGTTCTAGGCCCTG 34072
QY 1631 AGTAGAATGCTTAATCTTATTGACATTCAGTCTTTCTTCTTCTTCTTCTATGATGTTT 1690
Db 34073 AAATGCTTATAGGTTGTTTATTTGCAATCTGCTACTTTTATAAATAGAGCCATTTAT 34132
QY 1691 GCGATAAATTTCCCTTTAACTTCCCTTTTCCACTTCACTACATCTCACAAATTTGAT 1750
Db 34133 GCTATTAATTTTCTTTTAA-----AGACTACTTTTGTGATCTCTATAGTTGGGTA 34186
QY 1751 AGGAGTAGTTTAAATATCATTAGTATCTAAATATTTTAAATTTCTGTAT---TTTCTTC 1807

[illegible]

CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 16086 BP; 3675 A; 4642 C; 4404 G; 3365 T; 0 U; 0 Other;
 Query Match 9.4%; Score 239.4; DB 4; Length 16086;
 Best Local Similarity 70.8%; Pred. No. 2.5e-26;
 Matches 318; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
 QY 817 TGGTGTATATGACTATTATACGTCATATACAAATTTTATGCTATGATATATGACT 936
 DB 3055 TGTAAATGTTTGTGCTGATAGGCAATTTAGTTCACGTGCTGATTTATTATTATGAT 2996
 QY 877 ATTGTTCTCATATGATAAATTTTATGCTATGACAAATTTTATGCTTGTATATGACT 936
 DB 2995 ATATTGGGTTTATATCTCCCATCATCAATTTTGTCTATTTGCTCCACCTCTGTGTTCT 2936
 QY 937 TTATGTTATATGCTATACAACTAGATTTTGTGTTTGTGTTTGTACCGAGTCTTACTCTG 996
 DB 2935 TTTTCTCTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2876
 QY 997 TCACCCAGGCTGGAGTGAATGGCATGCTCTCAGCTCACTGCAACCTCCGCTCCCGGCT 1056
 DB 2875 TCACCCAGGCTGGAGTGAATGGCATGCTCTCAGCTCACTGCAACCTCCGCTCCCGGCT 2816
 QY 1057 TCAAGCCATTTCTCCACCTCAGCTCCCAAGTCTGGGACTCAGGATGAGCCACCGC 1116
 DB 2815 TCATGCCATCTCTGCTCCTCAGCTCCCAAGTCTGGGACTCAGGATGAGCCACCGC 2756
 QY 1117 ACCCGGCTAATTTTGTATTTTATAGTAGAGAGCTGGTTCCTATGTTGGCCAGGCTGAT 1176
 DB 2755 GGCCGGCTAATTTTGTATTTTATAGTAGAGAGCTGGTTCCTATGTTGGTTCAGGCTGTT 2696
 QY 1177 CTCGAATCTCTGACCTGTATCTCAGCCGCTCGGCTCCGCTCCCAAGTCTGGGATACAGG 1236
 DB 2695 CTCGAATCTCTGACCTGTATCTCAGCCGCTCGGCTCCGCTCCCAAGTCTGGGATACAGG 2636
 QY 1237 CGTGAGCCATTTGCTCTGGCCGATTTTCT 1265
 DB 2635 CGTGAGCCACCGGCTCGGCTTTTCT 2607
 RESULT 11
 AAL36020
 ID AAL36020 standard; DNA; 16086 BP.
 XX
 AC AAL36020;
 XX
 DT 08-JAN-2002 (first entry)
 XX
 DE Human musculoskeletal system related polynucleotide SEQ ID NO 2385.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein;
 KW musculoskeletal system; ds.
 XX
 OS Homo sapiens.
 XX
 FN WO200155367-A1.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001338.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.

16-MAR-2000; 2000US-0189874P.
 17-MAR-2000; 2000US-0190076P.
 18-APR-2000; 2000US-0198123P.
 19-MAY-2000; 2000US-0205515P.
 07-JUN-2000; 2000US-0209467P.
 28-JUN-2000; 2000US-0214886P.
 30-JUN-2000; 2000US-0215135P.
 07-JUL-2000; 2000US-0216647P.
 07-JUL-2000; 2000US-0216880P.
 11-JUL-2000; 2000US-0217487P.
 11-JUL-2000; 2000US-0217496P.
 14-JUL-2000; 2000US-0218290P.
 26-JUL-2000; 2000US-0220963P.
 26-JUL-2000; 2000US-0220964P.
 14-AUG-2000; 2000US-0224519P.
 14-AUG-2000; 2000US-0224519P.
 14-AUG-2000; 2000US-0225213P.
 14-AUG-2000; 2000US-0225214P.
 14-AUG-2000; 2000US-0225266P.
 14-AUG-2000; 2000US-0225267P.
 14-AUG-2000; 2000US-0225268P.
 14-AUG-2000; 2000US-0225270P.
 14-AUG-2000; 2000US-0225270P.
 14-AUG-2000; 2000US-0225447P.
 14-AUG-2000; 2000US-0225757P.
 14-AUG-2000; 2000US-0225758P.
 14-AUG-2000; 2000US-0225759P.
 18-AUG-2000; 2000US-0226279P.
 22-AUG-2000; 2000US-0226681P.
 22-AUG-2000; 2000US-0226686P.
 22-AUG-2000; 2000US-0227182P.
 23-AUG-2000; 2000US-0227009P.
 30-AUG-2000; 2000US-0228924P.
 01-SEP-2000; 2000US-0229287P.
 01-SEP-2000; 2000US-0229343P.
 01-SEP-2000; 2000US-0229344P.
 01-SEP-2000; 2000US-0229345P.
 05-SEP-2000; 2000US-0229509P.
 05-SEP-2000; 2000US-0229513P.
 06-SEP-2000; 2000US-0230437P.
 06-SEP-2000; 2000US-0230438P.
 08-SEP-2000; 2000US-0231242P.
 08-SEP-2000; 2000US-0231243P.
 08-SEP-2000; 2000US-0231244P.
 08-SEP-2000; 2000US-0231413P.
 08-SEP-2000; 2000US-0231414P.
 08-SEP-2000; 2000US-0232080P.
 08-SEP-2000; 2000US-0232081P.
 12-SEP-2000; 2000US-0232397P.
 14-SEP-2000; 2000US-0232397P.
 14-SEP-2000; 2000US-0232398P.
 14-SEP-2000; 2000US-0232399P.
 14-SEP-2000; 2000US-0232400P.
 14-SEP-2000; 2000US-0232401P.
 14-SEP-2000; 2000US-0233063P.
 14-SEP-2000; 2000US-0233064P.
 14-SEP-2000; 2000US-0233065P.
 21-SEP-2000; 2000US-0234223P.
 21-SEP-2000; 2000US-0234274P.
 25-SEP-2000; 2000US-0234997P.
 25-SEP-2000; 2000US-0234998P.
 26-SEP-2000; 2000US-0235484P.
 27-SEP-2000; 2000US-0235834P.
 27-SEP-2000; 2000US-0235836P.
 29-SEP-2000; 2000US-0236327P.
 29-SEP-2000; 2000US-0236367P.
 29-SEP-2000; 2000US-0236368P.
 29-SEP-2000; 2000US-0236369P.
 29-SEP-2000; 2000US-0236370P.
 02-OCT-2000; 2000US-0236802P.
 02-OCT-2000; 2000US-0237037P.
 02-OCT-2000; 2000US-0237038P.
 02-OCT-2000; 2000US-0237039P.
 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-02399335P.
PR 13-OCT-2000; 2000US-02399377P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPT; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
XX disorders related to the musculoskeletal system including musculoskeletal
XX cancers and also for testing and detection e.g. diagnosis.
XX
XX Example 2; SEQ ID NO 2385; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins

CC (ABB03087-ABB04109) associated with the musculoskeletal system useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. The genes are isolated from a range of human tissues
CC disclosed in the specification. The nucleic acids, proteins, antibodies
CC and (ant)agonists are useful in the diagnosis, treatment and prevention
CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, multiple sclerosis, rheumatoid
CC diabetes mellitus, Crohn's disease, (c) cardiovascular disorders such as
CC arthritis and ulcerative colitis; (d) wound healing; (e) neurological diseases e.g.
CC myocardial ischaemias; (f) and (g) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 16086 BP; 3365 A; 4404 C; 4642 G; 3675 T; 0 U; 0 Other;

Query Match 9.4%; Score 239.4; DB 4; Length 16086;
Best Local Similarity 70.8%; Pred. No. 2.5e-26;
Matches 318; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
QY 817 TGGTTGTAATACACTATTTTACGTCATATACAAATTTTATGTTTGTAAATGACTTT 876
DB 13032 TGTAAATGTTTGAAGTGTATAGGCGCAATTTAGTCTCAATGCTGATTTTATATGGAT 13091
QY 877 ATGTTTGTCAATGATAATTTTATGTCATAGAACAAATTTTATGCTTGATATGACT 936
DB 13092 ATATTGGGTTATATCTCCCATCACATTTTGTGCTATTGTCACCTCTTGTTCT 13151
QY 937 TTATGTTTATAGGCTATACAACTAGATTTTTTGTGTTTGTGACCGAGCTTACTCTG 996
DB 13152 TTCTCTCCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 13211
QY 997 TCACCCAGGCTGAGTGTAAATGGCATGTTCTCAGCTCAGTGCACCTCCGCTCCCGGT 1056
DB 13212 TCACCCAGGCTGAGTGTAAATGGCATGTTCTCAGCTCAGTGCACCTCCGCTCCCGGT 13271
QY 1057 TCAAGCCATTCTTCCACCTCAGCTCCCAAGTAGTGGGACTACAGGCATGAGCCACCG 1116
DB 13272 TCATGCCATCTCTGCTCAGCTCCCAAGTAGTGGGACTACAGGCATGAGCCACCG 13331
QY 1117 ACCCGGCTAATTTTGTATTTTATAGTACAGCTGTTTCCATATGTTGCCAGGCTGAT 1176
DB 13332 GGCCTGCTAATTTTGTATTTTATAGTACAGCTGTTTCCATATGTTGCCAGGCTGAT 13391
QY 1177 CTGAACTCTGACCTGTAAATCCACCGCTCGGCTCCCAAGTCTCGGATTCAGG 1236
DB 13392 CTCGAACTCTGACCTGTGATCTGCGCTCTGCGCTCCCAAGTCTCGGATTCAGG 13451
QY 1237 CGTGAGCCATTGCTGCTGGCGGATTTTCT 1265
DB 13452 CGTGAGCCCGGCTGCGCTTTTCT 13480

RESULT 12
ABX59008
ID ABX59008 standard; cDNA; 16086 BP.
XX
XX AC ABX59008;
XX
XX DT 26-FEB-2003 (first entry)
XX
XX DE cDNA encoding novel human musculoskeletal system antigen #1352.
XX
XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
XX post-operative tissue repair; limb regeneration; neuronal growth;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX AIDS-related complex; chondrocyte growth; bone regeneration;

Db 13332 GCGCGCTAATATTATTTTATTTAGTAGAGACGGGTTTACCAATGTTGTCAGGCTGGT 13391
QY 1177 CTGCAACTCTGACCTTGTAATCCACCGGCTCGGCTGCAAAAGTCTGGGATTACAGG 1236
|||||
Db 13392 CTGCAACTCTGACCTTGTAATCTGCGGCTTGGCTCCAAAGTCTGGGATTACAGG 13451
QY 1237 CGTGAGCCATGTCCTGCGCGGATTTTT 1265
|||||
Db 13452 CGTGAGCCACCGCGCTGGCCTTTTCT 13480

RESULT 13
AAL04865/c
ID AAL04865 standard; DNA; 538 BP.
XX AAL04865;
XX AC
XX AC
DT 21-NOV-2001 (first entry)
XX
XX Human reproductive system related antigen DNA SEQ ID NO: 7553.
DE Human reproductive system related antigen; reproductive system disorder;
KW Human; reproductive system related antigen; cancer; gene therapy; ds.
KW cancer; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX WC200155320-A2.
PN
XX
XX
PD 02-AUG-2001.
XX
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186630P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; SEQ ID NO 7553; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX
XX Sequence 538 BP; 151 A; 134 C; 139 G; 114 T; 0 U; 0 Other;
SQ
Query Match 9.4%; Score 238.8; DB 4; Length 538;
Best Local Similarity 89.0%; Pred. NO. 4e-26;
Matches 258; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
Qy 967 TTTTGTGTTTTCACCGAGTCTTACTCTGTACCCAGGCTGGAGTGAATGGCATGTC 1026
Db 533 TTTTGTGTTTTCACAGAGTCTACTCTGTACCCAGGCTGGAGTGAATGGCATGTC 474
Qy 1027 TCAGCTCACTGCAACCTCCGCTCCCGGGTTCAAGCCATTCTCCACCTCAGCCTCCCAA 1086
Db 473 TCCACTCACTGCAACCTCCACCTCCCGGGTTCAAGCGATTTCACCTCAGCCTCCCGA 414
Qy 1087 GTAGCTGGGACTACAGGATGAGCAGCCGACCGCGCTAATTTTGTATTTTAGTAGAG 1146
Db 413 GTAGCTGGGACTACAGGATGAGCAGCCGACCGCGCTGATTTTGTATTTTAGTAGAG 354
Qy 1147 ACGTGGTTCCACTATGTTGGCCAGCTGATCTCGAACTCCTGACCTTCTATCCACCGC 1206
Db 353 ATGGGGTTTCACTAATGTTGGCCAGCTGCTCTAACTCTCTGACCTCATGATCCACCTGC 294
Qy 1207 CTCGGCTGCCAAAGTGTGGGATTACAGGGGTGAGCGTATGTGCTGGC 1256
Db 293 CTCGGCTGCCAAAGTGTGGGATTACAGGGGTGAGCGTATGTGCTGGC 244
RESULT 14
AAL04867/c
ID AAL04867 standard; DNA; 538 BP.
XX
XX AAL04867;
XX

DT 21-NOV-2001 (first entry)
XX Human reproductive system related antigen DNA SEQ ID NO: 7555.
DE Human; reproductive system related antigen; reproductive system disorder;
XX cancer; gene therapy; ds.
KW Homo sapiens.
XX WO200155320-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001339.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.

14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 14-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 21-SEP-2000; 2000US-0234977P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239353P.
 PR 13-OCT-2000; 2000US-0239357P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249284P.
 PR 17-NOV-2000; 2000US-0249285P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465570/50.
 DR
 XX Isolated nucleic acid molecule encoding a reproductive system antigen is
 PT used in preventing, treating or ameliorating a medical condition.
 XX Disclosure; SEQ ID NO 7555; 1297pp + Sequence Listing; English.
 XX The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a genomic sequence encoding a
 CC protein of the invention
 XX
 SQ Sequence 538 BP; 151 A; 134 C; 139 G; 114 T; 0 U; 0 Other;
 Query Match 9.4%; Score 238.8; DB 4; Length 538;
 Best Local Similarity 89.0%; Pred. No. 4e-26;
 Matches 258; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 QY 967 TTTTGTGTTTGGACCGAGTCTTACTCTCTACCCAGGCTGGAGTGAATGGCATGGTC 1026
 DB 533 TTTTGTGTTTGGACCGAGTCTTACTCTCTACCCAGGCTGGAGTGAATGGCATGGTC 474
 QY 1027 TCAGCTCACTGCAACCTCCGCTCCCGGGTTCAAGCAATCTTCCACCTCAGCCTCCCA 1086
 DB 473 TCCACTCACTGCAACCTCCGCTCCCGGGTTCAAGCAATCTTCCACCTCAGCCTCCCA 414
 QY 1087 GTAGCTGGGACTACAGGCATGACCGACCGGCTGAATTTTGTATTTTAGTAGAG 1146
 DB 413 GTAGCTGGGACTACAGGCATGACCGACCGGCTGAATTTTGTATTTTAGTAGAG 354
 QY 1147 ACCTGGTTCACATATGTTGGCCAGGCTGATCTGAACTCTGACCTTGTATTCACCCGC 1206
 DB 353 ATGGGGTTTCACTATGTTGGCCAGGCTGATCTGAACTCTGACCTTGTATTCACCCGC 294
 QY 1207 CTCGGCTGCCAAAGTCTGGGATTACAGCGGTGAGCAATGTGCTCTGGC 1256
 DB 293 CTCGGCTGCCAAAGTCTGGGATTACAGCGGTGAGCAATGTGCTCTGGC 244
 RESULT 15
 ABL97761/c
 ID ABL97761 standard; DNA; 538 BP.
 XX
 AC ABL97761;
 XX
 DT 21-JUN-2002 (first entry)
 XX
 DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2413.
 XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;
 KW reproductive system disorder; urinary system disorder; gene therapy;
 KW cardiovascular disorder; respiratory system disorder; neurological disorder;
 KW gastrointestinal disease; infection; cytostatic; gene; ds.
 XX Homo sapiens.
 OS
 XX
 PN W0200155317-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001329.
 XX

PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198121P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216800P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249266P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-025179P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483232/52.

Nucleic acids encoding 973 human testicular antigen polypeptides, useful for preventing, diagnosing and/or treating testicular cancer.

XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 16:26:10 ; Search time 10196.7 Seconds
(without alignments)
10801.019 Million cell updates/sec

Title: US-09-914-651A-2_COPY_14040_16580
Perfect score: 2541
Sequence: 1 agcataatttttttttaatt.....cttctctagtgtgaatttt 2541

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2541	100.0	93525	9	AF007544	Homo sapi
2	2510	98.8	129095	9	AC110742	Homo sapi
3	2508.4	98.7	158524	2	AL162372	Homo sapi
c	2508.4	98.7	187529	9	AC118273	Homo sapi
4	2507.4	93.2	135637	9	AF003400	Homo sapi
5	2367	93.2	137888	9	AP003122	Homo sapi
6	2367	93.2	192648	2	AC024234	Homo sapi
c	2367	93.2	192648	2	AC024234	Homo sapi
7	2148	84.5	117521	2	AC023784	Homo sapi
c	2148	84.5	117521	2	AC023784	Homo sapi
8	1476.6	58.1	246865	2	AF040003	Homo sapi
9	1461.4	57.5	1669	9	AF480875	Homo sapi
10	930.8	36.6	246865	2	AC074003	Homo sapi
11	653	25.7	68842	2	AC137087	Homo sapi
c	653	25.7	68842	2	AC137087	Homo sapi
12	603.8	23.8	72409	2	AC131257	Homo sapi
c	603.8	23.8	72409	2	AC131257	Homo sapi
13	346.2	13.6	40131	2	AC136711	Homo sapi
14	335	13.2	335	6	BD269371	Controlli
15	321.2	12.6	157342	9	AC108696	Homo sapi
16	321.2	12.6	175233	2	AC022941	Homo sapi
17	320.2	12.6	49838	9	HS387E22	Human DNA
18	312	12.3	192802	9	AC005284	Homo sapi
19	308.6	12.1	141742	9	AC092621	Homo sapi
c	308.6	12.1	141742	9	AC092621	Homo sapi
20	298	11.7	87097	9	AC109344	Homo sapi
21	298	11.7	169261	2	AC140480	Homo sapi
22	296.8	11.7	156957	9	AP001977	Homo sapi
23	296.8	11.7	171752	2	AC147307	Homo sapi
c	296.8	11.7	171752	2	AC147307	Homo sapi
24	294.4	11.5	98842	9	AC008919	Homo sapi
25	293.4	11.5	108162	2	AC020960	Mus muscu
c	293.4	11.5	108162	2	AC020960	Mus muscu
26	293.4	11.5	124235	9	AC020893	Homo sapi
c	293.4	11.5	124235	9	AC020893	Homo sapi
27	293.4	11.5	152905	2	AC044801	Homo sapi
c	293.4	11.5	152905	2	AC044801	Homo sapi
28	292.8	11.5	105211	9	AC110076	Homo sapi
29	292.6	11.5	125402	9	AC114815	Homo sapi
30	291.6	11.5	83248	9	AC068754	Homo sapi
c	291.6	11.5	83248	9	AC068754	Homo sapi
31	291.6	11.5	173724	2	AC026766	Homo sapi
c	291.6	11.5	173724	2	AC026766	Homo sapi
32	291	11.5	172705	9	AC013272	Homo sapi
33	290.2	11.4	78794	9	AC003990	Homo sapi
c	290.2	11.4	78794	9	AC003990	Homo sapi
34	287.4	11.3	41572	9	AY217547	Homo sapi
35	287.4	11.3	159446	9	AC003669	Homo sapi
c	287.4	11.3	159446	9	AC003669	Homo sapi
36	287.4	11.3	190752	2	AC114805	Homo sapi
37	287.4	11.3	194492	9	AC097625	Homo sapi
38	286	11.3	109601	9	AC074392	Homo sapi
39	284.6	11.2	170940	9	AC120042	Homo sapi
40	284.6	11.2	181694	9	AC019357	Homo sapi
c	284.6	11.2	181694	9	AC019357	Homo sapi
41	282.2	11.1	187883	9	AC008268	Homo sapi

ALIGNMENTS

RESULT 1
AF007544
LOCUS Homo sapiens prostate-specific membrane antigen (PSM) gene, 93525 bp DNA linear PRI 10-DEC-1998
DEFINITION complete cds.
ACCESSION AF007544
VERSION 1 (bases 1 to 93525)
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS O'Keefe,D.S., Su,S.L., Bacich,D.J., Horiguchi,Y., Luo,Y., Powell,C.T., Zandvliet,D., Russell,P.J., Molloy,P.L., Nowak,N.J.,

TITLE	Shows, T.B., Mullins, C., Vonder Haar, R.A., Fair, W.R. and Heston, W.D. Mapping, genomic organization and promoter analysis of the human prostate-specific membrane antigen gene
JOURNAL	Biochim. Biophys. Acta 1443 (1-2), 113-127 (1998)
MEDLINE	99057588
PUBMED	9938072
REFERENCE	2 (bases 1682 to 65315)
AUTHORS	Heston, W.D.W., Su, S.L., Luo, Y., Huryk, R., Bacich, D.J., Fair, W.R., Mullins, C. and Vonder Haar, R.A.
TITLE	Direct Submission
JOURNAL	Submitted (06-JUN-1997) Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA
REFERENCE	3 (bases 1 to 93525)
AUTHORS	Powell, C.T., Sandvliet, D., Russell, P.J., Molloy, P.L., Nowak, N.J., Mullins, C., Vonder Haar, R.A., Fair, W.R. and Heston, W.D.W.
TITLE	Direct Submission
JOURNAL	Submitted (18-MAR-1998) Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA
REMARK	Sequence update by submitter
COMMENT	On Mar 19, 1998 this sequence version replaced gi:2897945.
FEATURES	Location/Qualifiers
source	1..93525
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="11"
	/map="11p11-12"
	2488..64518
gene	/gene="PSM"
mRNA	join(2488..2863,4994..5099,10726..10912,18275..18376,24400..24525,25314..25500,27927..28020,35216..35314,36196..36281,37697..37816,39896..39978,41911..41974,46402..46469,53129..53220,54364..54454,56661..56925,57226..57307,62423..62515,64209..64518)
	/product="prostate-specific membrane antigen"
CDS	join(2746..2863,4994..5099,10726..10912,18275..18376,24400..24525,25314..25500,27927..28020,35216..35314,36196..36281,37697..37816,39896..39978,41911..41974,46402..46469,53129..53220,54364..54454,56661..56925,57226..57307,62423..62515,64209..64398)
	/gene="PSM"
	/note="PSMA"
	/codon_start=1
	/product="prostate-specific membrane antigen"
	/protein_id="AAC83972.1"
	/db_xref="GI:2897946"
	/translation="MMWLLHETDSAVATARRPRLWCAGLVLAGGPFLLGLFGWFIKSSNEATNITPKHNMKAFDLDELKAEINKFLHFTQIPHLAGTEQNFOLAKOIQSQWKFGLDVSELAHVDLLSPKNTHPNYSIINEDGNEIFNTSLFPPPEGYENVSDIYPPFSAPSGQMEGLVYVYARTEDFKLERDMKINCISKVIARYKVFGRGNKVNQAQYAKAGVILSDPADYFAPGVKSPDGWNLPGGVORGNILNLNGAGDPLTPGVPALEYARGAEAVGLSPFVPHFGIYDAQKLEKMGSGAPPSSWRGSLKVPYNYGPGTGNFTQKQHHIHNTEVRIYNYVIGLRGAVPDRVILGGRHDSWVFGIDPQSGANVHEIVRSFGTILKEGWRRPRTILFASDAEFGLLGSTWAEENSRDLQERGVAYINADSIKGNVTLRVDCTPLMYSLVHNLTKELKSPDEFGKSLYESWTKSPSPFESGMPTSLKLGSDNRPVFFORGLIAGRRARYTKNWTNKFSGVPLVSHVYETVELVEKEYDPEKHLTVAQVGRGVFLANSIVLPFCRDYAVLRKVADKYSISMKHPQEMKTYSVSFDLSIFRAKNTFEIASKPSERLQDFKSNPILVRMNDQLMFLERAFIDPLGRDRPFYRHVIYAPSSHNNKYGESFFGIYDALFDIESKVDPSKANGVEVAKQIVAAFTVQAAAEITLSEVA"
ORIGIN	
Query Match	100.0%; Score 2541; DB 9; Length 93525;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2541; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 AGCATATATTTTCTTTTAACTTCTACCTAGTCAATTCAGATGAGTTCAATTTCTCATGTGAAT 60
Db	14040 AGCATATATTTTCTTTTAACTTCTACCTAGTCAATTCAGATGAGTTCAATTTCTCATGTGAAT 14099
QY	61 CACAGAAAAAATATGGAAATTTGGAATATGTGGGGATGTGGCAAGTCGTAGTTGATT 120

Db	14100 CACAGAAAAAATATGGAAATTTGGAATATGTGGGGATGTGGCAAGTCGTAGTTGATT 14159
QY	121 GGTTCATTTTACCTTCTATCCATGTGAGAAAAGGAAATAACTACTGCTCTAGTCAGTA 180
Db	14160 GGTTCATTTTACCTTCTATCCATGTGAGAAAAGGAAATAACTACTGCTCTAGTCAGTA 14219
QY	181 CAATGTAAATCTGCCAGATCCCTCTCATCTCCAAATCCCTCCCAATCTGTAATCTG 240
Db	14220 CAATGTAAATCTGCCAGATCCCTCTCATCTCCAAATCCCTCCCAATCTGTAATCTG 14279
QY	241 CACTGACCAAGACCTATACAAATATATGTTGAAAAGAGTTTGTAGACAGATATCCTTA 300
Db	14280 CACTGACCAAGACCTATACAAATATATGTTGAAAAGAGTTTGTAGACAGATATCCTTA 14339
QY	301 ACTTTCTCTGGTCTTTGAAAGAAATACCTTTAATCTTCTTCCACCAATTAATTAATG 360
Db	14340 ACTTTCTCTGGTCTTTGAAAGAAATACCTTTAATCTTCTTCCACCAATTAATTAATG 14399
QY	361 TACTGTTTCTTTTCTGTTATCTTAAAGTACATCTTCCACATATATCTAAATTTT 420
Db	14400 TACTGTTTCTTTTCTGTTATCTTAAAGTACATCTTCCACATATATCTAAATTTT 14459
QY	421 TGCATGATGATTTATATTAACACATTTTCTTAAAGTACATCTTCCACATATATCTTCTTC 480
Db	14460 TGCATGATGATTTATATTAACACATTTTCTTAAAGTACATCTTCCACATATATCTTCTTC 14519
QY	481 TTGATCTGAAGATGACATTCAGTGATATCTAATGTTAGAAATCTTATATTTTGGAA 540
Db	14520 TTGATCTGAAGATGACATTCAGTGATATCTAATGTTAGAAATCTTATATTTTGGAA 14579
QY	541 CTGTGATATATAGCATTAGGTTTAAATTCATATATTCATATATATATATGTTTATCTC 600
Db	14580 CTGTGATATATAGCATTAGGTTTAAATTCATATATTCATATATATATATGTTTATCTC 14639
QY	601 ATAGCATGTTTACTGTTGTTTGGCAATCTATTTGTTGGTTTAAATTTCAATATTTCTA 660
Db	14640 ATAGCATGTTTACTGTTGTTTGGCAATCTATTTGTTGGTTTAAATTTCAATATTTCTA 14699
QY	661 CTTCGCTTCTAAATAGTGGTTTAAATTTCTTGAAGTAGTTTATTTGCAATTA 720
Db	14700 CTTCGCTTCTAAATAGTGGTTTAAATTTCTTGAAGTAGTTTATTTGCAATTA 14759
QY	721 AATATATTTTCTTTAACTTCAACTCAAGGAAACAGTTGGCTTGGCTTGGCTTGGCTT 780
Db	14760 AATATATTTTCTTTAACTTCAACTCAAGGAAACAGTTGGCTTGGCTTGGCTTGGCTT 14819
QY	781 GTGAAAAATTTTAACTTCAACTTCTTTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
Db	14820 GTGAAAAATTTTAACTTCAACTTCTTTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 14879
QY	841 TCATATAACAATTTTATTTGTTTAAATGACCTTATTTGTTGTTGTTGTTGTTGTTGTTGTT 900
Db	14880 TCATATAACAATTTTATTTGTTTAAATGACCTTATTTGTTGTTGTTGTTGTTGTTGTTGTT 14939
QY	901 TGTATAGAACAAATTTTATTTGTTGTTATATGCTTATTTGTTGTTGTTGTTGTTGTTGTTGTT 960
Db	14940 TGTATAGAACAAATTTTATTTGTTGTTATATGCTTATTTGTTGTTGTTGTTGTTGTTGTTGTT 14999
QY	961 AGATTTTGTGTTGTTTTCACCGAGTCTTCTCTGTCACCGAGCTGGAGTGTAAATGTC 1020
Db	15000 AGATTTTGTGTTGTTTTCACCGAGTCTTCTCTGTCACCGAGCTGGAGTGTAAATGTC 15059
QY	1021 ATGCTCTCAGCTCACTCAACCTCCGCTCCCGGTTCAAGCCATTTCTCCACCTCAGCC 1080
Db	15060 ATGCTCTCAGCTCACTCAACCTCCGCTCCCGGTTCAAGCCATTTCTCCACCTCAGCC 15119
QY	1081 TCCCAAGTAGCTGGAGTACAGGATGAGCCACCGCCCGCTTAAATTTTGTATTTTATTTA 1140
Db	15120 TCCCAAGTAGCTGGAGTACAGGATGAGCCACCGCCCGCTTAAATTTTGTATTTTATTTA 15179
QY	1141 GTAGAGACGTGGTCCCATATATGTTGGCAGGCTGATCTCGAACTCTCGAACTCTGAAATCC 1200

Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Minova, T., Miengo, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (27-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 129095)

Biren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Minova, T., Miengo, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (02-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 2, 2002 this sequence version replaced gi:20334649.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submission@genome.wi.mit.edu

Project Information
Center project name: L21820
Center clone name: 2026_G_22

Location/Qualifiers
1. .129095
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="CTD-2026G22"
/clone_lib="CITD1 Human BAC"
2. .261
/rpt_family="L1PA6"
1334..1562
/rpt_family="MIR"
complement(2716..2929)
/rpt_family="L2"
3055..3336
/rpt_family="AluJo"
2468..3524

repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

/rpt_family="(TTCC)n"
3525..3560
/rpt_family="(TTCC)n"
4640..4849
/rpt_family="MER20"
complement(4884..5129)
/rpt_family="L3"
complement(5208..5413)
/rpt_family="MIR"
5840..5891
/rpt_family="L1MC5"
complement(6065..8047)
/rpt_family="L1PA6"
8050..10029
/rpt_family="L1PA6"
10340..10497
/rpt_family="L1ME4A"
10670..10691
/rpt_family="AT_rich"
10977..10997
/rpt_family="AT_rich"
complement(10998..11293)
/rpt_family="AluJo"
12004..12042
/rpt_family="AT_rich"
12337..12396
/rpt_family="(TA)n"
12447..12489
/rpt_family="AT_rich"
12950..12970
/rpt_family="AT_rich"
complement(13047..13097)
/rpt_family="MIR3"
13218..13339
/rpt_family="MIR"
13460..13500
/rpt_family="AT_rich"
14066..14087
/rpt_family="AT_rich"
14141..14166
/rpt_family="(TTTTTG)n"
15042..15277
/rpt_family="AluSp"
complement(16069..16260)
/rpt_family="MIR"
complement(16362..16449)
/rpt_family="MIR"
16603..16637
/rpt_family="(TG)n"
16976..17065
/rpt_family="L1"
17284..17286
/rpt_family="(A)n"
17817..17860
/rpt_family="(TTTG)n"
18317..18432
/rpt_family="L2"
18922..19089
/rpt_family="MER45A"
20139..20250
/rpt_family="FLAM_C"
20252..20292
/rpt_family="(TAAAA)n"
20804..21817
/rpt_family="L1MA3"
21815..22853
/rpt_family="L1MA3"
22833..24145
/rpt_family="L1MA3"
24826..24896
/rpt_family="AT_rich"
25215..25380
/rpt_family="MIR"

repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

repeat_region	25381..25402	/rpt_family=" (TG) n"
repeat_region	25403..25440	/rpt_family="MIR"
repeat_region	complement(26317..26407)	
Query Match	98.8%; Score 2510; DB 9; Length 129095;	
Best Local Similarity	99.7%; Pred. No. 0;	
Matches 2535; Conservative	0; Mismatches 5; Indels 2; Gaps 2;	
QY	1 AGCATAATTTTCTTTTAACTCTTACTAGTGGAATCAGATGAGTTCATTTTCTCATGTGAAT	60
Db	122651 AGCATPAATTTTCTTTTAACTCTTACTAGTGGAATCAGATGAGTTCATTTTCTCATGTGAAT	122710
QY	61 CACAGAAAAAATATGGGAAATTTGGAATATGTGGGATGTGCGAAGTCGTAGTTGATTT	120
Db	122711 CACAGAAAAAATATGGGAAATTTGGAATATGTGGGATGTGCGAAGTCGTAGTTGATTT	122770
QY	121 GGTTCATTTTGTCTTCTATCCATGTGAGAAAGGGAATAACTACTGTCTAGTCAGTA	180
Db	122771 GGTTCATTTTGTCTTCTATCCATGTGAGAAAGGGAATAACTACTGTCTAGTCAGTA	122830
QY	181 CAATCTAAATCTGCCAGATCCCTCCTCATCTCCAAATCCCTCCCATTTCTGATTTACTG	240
Db	122831 CAATCTAAATCTGCCAGATCCCTCCTCATCTCCAAATCCCTCCCATTTCTGATTTACTG	122890
QY	241 CACTGACCAAGACCTTATACAATATTAATGTGAAAAGAGTTTGTATAGCAGATATCCCTTA	300
Db	122891 CACTGACCAAGACCTTATACAATATTAATGTGAAAAGAGTTTGTATAGCAGATATCCCTTA	122950
QY	301 ACTTCTCTCGTCTTGAAGGAATACTTTAAATCTTTTCTCACCATTAAATATGATTT	360
Db	122951 ACTTCTCTCGTCTTGAAGGAATACTTTAAATCTTTTCTCACCATTAAATATGATTT	123010
QY	361 TACTGTTTTCTTTTCTCGTTTCTTAAAGTACATCCCTTCCACATATACATAAAATTTT	420
Db	123011 TACTGTTTTCTTTTCTCGTTTCTTAAAGTACATCCCTTCCACATATACATAAAATTTT	123070
QY	421 TGCTGATGGATTTATATTAACAACATTTTCTTTAAGATGAGCATGGATTTTCTTTC	480
Db	123071 TGCTGATGGATTTATATTAACAACATTTTCTTTAAGATGAGCATGGATTTTCTTTC	123130
QY	481 TTGATCTGAAGATGACATCTCAGTGATATCTTAATGTTAGAAATTAATCTTATATTTTGA	540
Db	123131 TTGATCTGAAGATGACATCTCAGTG-TAATCTAAATGTTAGAAATTAATCTTATATTTTGA	123189
QY	541 CTTCTAATATTAGCATATAGGTTTTAAATTCATATATTGATATATATTATGTTTATCTC	600
Db	123190 CTTCTAATATTAGCATATAGGTTTTAAATTCATATATTGATATATATTATGTTTATCTC	123249
QY	601 ATAGCATGTTTGTGCTGTTTTTGCAATCTAATTTGTTGGGTATAAATTCGAATTTCTTA	660
Db	123250 ATAGCATGTTGTGCTGTTTTTGCAATCTAATTTGTTGGGTATAAATTCGAATTTCTTA	123309
QY	661 CTTGCCCTTAAATAGAGTTCGGTTTTTAAATTTTCTGAAGTAGGTTTTTATGCAATTA	720
Db	123310 CTTGCCCTTAAATAGAGTTCGGTTTTTAAATTTTCTGAAGTAGGTTTTTATGCAATTA	123369
QY	721 AATTAATTTTTCCTTTAACTTCAAACCTCAAGGAAACCCAGTTGGCCCTTGACTCTGTTT	780
Db	123370 AATTAATTTTTCCTTTAACTTCAAACCTCAAGGAAACCCAGTTGGCCCTTGACTCTGTTT	123429
QY	781 GTGGAAAAATTTTAAACTACTGTTTTAAATTTCTTTATTCGTTCTAATATGACATTTTACG	840
Db	123430 GTGGAAAAATTTTAAACTACTGTTTTAAATTTCTTTATTCGTTCTAATATGACATTTTACG	123489
QY	841 TCATATAACAATTTTATTTGTTTGAATGACCTTTAATTTGTTGTCATATGATTAATTTTA	900
Db	123490 TCATATAACAATTTTATTTGTTTGAATGACCTTTAATTTGTTGTCATATGATTAATTTTA	123549
QY	901 TGTCAATAGAACAAATTTTATTTGCTTGATATATGACTTTTATTTGTTATATGGCTATACAACT	960
Db	123550 TGTCAATAGAACAAATTTTATTTGCTTGATATATGACTTTTATTTGTTATATGGCTATACAACT	123609

```
QY 2040 TTTAATATTTTGTATTTGATTTGACATTTAGTTTGGCTTCTTATTTAGTTTGGCTGAAATCTCCA 2099
|
|
|
Db 124690 TTTAATATTTTGTATTTGATTTGACATTTAGTTTGGCTTCTTATTTAGTTTGGCTGAAATCTCCA 124749
|
|
|
QY 2100 TTATCCTTAAATGCTCTCTCATTTTGTCTGCTTCTTATTTAATTTAGAGATAAAATGTTA 2159
|
|
|
Db 124750 TTATCCTTAAATGCTCTCTCATTTTGTCTGCTTCTTATTTAATTTAGAGATAAAATGTTA 124809
|
|
|
QY 2160 AATTATCTCACCCTCAGTATAGTATGCTCTGTTTATATATATATATATAAATTTATAAT 2219
|
|
|
Db 124810 AATTATCTCACCCTCAGTATAGTATGCTCTGTTTATATATATATATAAATTTATAAT 124869
|
|
|
QY 2220 CCATAAATTTATGTTATGTTATTTTGGAGACCTTATATCATATATAAAGCAATTTGTTG 2279
|
|
|
Db 124870 CCATAAATTTATGTTATGTTATTTTGGAGACCTTATATCATATATAAAGCAATTTGTTG 124929
|
|
|
QY 2280 ATGAATGACAGACTTATATCTATGTTAGTATGCTTTTATCTCGTCATATAATGTTATTG 2339
|
|
|
Db 124930 ATGAATGACAGACTTATATCTATGTTAGTATGCTTTTATCTCGTCATATAATGTTATTG 124989
|
|
|
QY 2340 ACTTTGCTCTAAATTTTATTAATTAATTTTGTGTTATTTTCTTTTTCAGCGGGTTT 2399
|
|
|
Db 124990 ACTTTGCTCTAAATTTTATTAATTAATTTTGTGTTATTTTCTTTTTCAGCGGGTTT 125049
|
|
|
QY 2400 ATGTACTGCTGTCAATTTGTTACACAGCTGATTTTATTTAGACATGCTACGCTTTTAA 2459
|
|
|
Db 125050 ATGTACTGCTGTCAATTTGTTACACAGCTGATTTTATTTAGACATGCTACGCTTTTAA 125109
|
|
|
QY 2460 TTATTTCTTTTTCATTTTCAATTTTATTAATTTCTGATATACATAATTTAGTCACTTTT 2519
|
|
|
Db 125110 TTATTTCTTTTTCATTTTCAATTTTATTAATTTCTGATATACATAATTTAGTCACTTTT 125169
|
|
|
QY 2520 ACCTTCCTCTAGTGTGAATTTT 2541
|
|
|
Db 125170 ACCTTCCTCTAGTGTGAATTTT 125191
|
|
|
RESULT 3
AL162372 158524 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome 13 clone RP11-181P16, 3 unordered pieces.
DEFINITION AL162372
ACCESSION AL162372.11 GI:12539555
VERSION HTG; HTGS PHASE1; HTGS_CANCELLED.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Burton, J.
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jan 25, 2001 this sequence version replaced gi:11693478.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA181P16
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 157639 bases at least Q40
Consensus quality: 157977 bases at least Q30
Consensus quality: 158153 bases at least Q20
Insert size: 158324; sum-of-contigs
Insert size: 174362; 5.7% error; agarose-fp
Quality coverage: 6.90x in Q20 bases; agarose-fp
coverage: 6.26x in Q20 bases; agarose-fp

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 59356: contig of 59356 bp in length
* 59357 59456: gap of 100 bp
* 59457 65248: contig of 5792 bp in length
* 65249 65349: gap of 100 bp
* 65349 158524: contig of 93176 bp in length.
FEATURES
Location/Qualifiers
1..158524
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="13"
/clone_lib="RP11-181P16"
/clone_lib="RP11-181P16"
1..59356
/notes="assembly_fragment:00306
clone_end:17
vector_side:left"
59457..65248
/notes="assembly_fragment:01282
fragment_chain:1"
65349..158524
/notes="assembly_fragment:02815
fragment_chain:1"
misc_feature
misc_feature
misc_feature
misc_feature
ORIGIN
Query Match 98.7%; Score 2508.4; DB 2; Length 158524;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2534; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 1 AGCATATTTTCTTTTAACTTTACCTAGTAGTACAGATGAGTTCAATTTCTCATGTGAAT 60
Db 70774 AGCATATTTTCTTTTAACTTTTACCTAGTAGTACAGATGAGTTCAATTTCTCATGTGAAT 70833
QY 61 CACAGAAAAATATGCGGAAATTTGGAATATGTTGGGAGTGTGCAAGTGTGAGTTGATTT 120
Db 70834 CACAGAAAAATATGCGGAAATTTGGAATATGTTGGGAGTGTGCAAGTGTGAGTTGATTT 70893
QY 121 GGTTCATTTTACCTTCTATCCATGTCAGAAAAAGGAAATTAATCTGCTCTAGTCAGTA 180
Db 70894 GGTTCATTTTACCTTCTATCCATGTCAGAAAAAGGAAATTAATCTGCTCTAGTCAGTA 70953
QY 181 CAATGTAAATCTGCCAGATCCCTCTCATCTCCAAATCCCTCCCAATTTGTTATCTG 240
Db 70954 CAATGTAAATCTGCCAGATCCCTCTCATCTCCAAATCCCTCCCAATTTGTTATCTG 71013
QY 241 CACTGACCAAGCCTATACAAATATATGTTGAAAGAGTTTGTATGAGCAGATATCCTTA 300
Db 71014 CACTGACCAAGCCTATACAAATATATGTTGAAAGAGTTTGTATGAGCAGATATCCTTA 71073
QY 301 ACTTTCTTCTGGTCTTTGAAGAAATACCTTTTAAATCTTCTCCACCAATTAATTTGTTAT 360
Db 71074 ACTTTCTTCTGGTCTTTGAAGAAATACCTTTTAAATCTTCTCCACCAATTAATTTGTTAT 71133
QY 361 TACTGTTTCTTTTCTGGTTATCTTAAAGTACATCTTTTCCACATATACATAAATTTT 420
Db 71134 TACTGTTTCTTTTCTGGTTATCTTAAAGTACATCTTTTCCACATATACATAAATTTT 71193
QY 421 TGCTGTATGCAATTTATTAACACATATTTTCTTTTAAGATGAGCAGTCGATTTTCTTTC 480
Db 71194 TGCTGTATGCAATTTATTAACACATATTTTCTTTTAAGATGAGCAGTCGATTTTCTTTC 71253
QY 481 TTGATCTGAAGATGACATCTCAGTGATTTCTAATGTTAGTAATCTTATTTATTTTGA 540
Db 71254 TTGATCTGAAGATGACATCTCAGTG- TATTTCTAATGTTAGTAATCTTATTTATTTTGA 71312
```


ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 187529)

REFERENCE

AUTHORS

Birren, B., Nusbaum, C. and Lander, E.

TITLE

Homo sapiens chromosome 11, clone RP11-107P7

JOURNAL

Unpublished

REFERENCE

AUTHORS

2 (bases 1 to 187529)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R., Landers, T., Lehotsky, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE

Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

AUTHORS

3 (bases 1 to 187529)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE

Submitted (29-AUG-2002) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

AUTHORS

4 (bases 1 to 187529)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (09-SEP-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 9, 2002 this sequence version replaced gi:22539138.

All repeats were identified using RepeatMasker:

Smit, A.P.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submission@genome.wi.mit.edu

----- Project Information

Center project name: L26235

Center clone name: 107_P_7

Location/Qualifiers

1. 187529

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="RP11-107P7"

/clone_lib="RP11-107P7"

90. 113

/rpt_family="AT_rich"

1137. 1285

/rpt_family="MLT1E3"

1445. 1561

/rpt_family="MLT2F"

2069. 2350

/rpt_family="AluJo"

3061. 3100

/rpt_family="(TATATG)n"

complement(3537. 3585)

/rpt_family="MLT2B1"

3608. 3880

/rpt_family="MER7A"

3881. 4194

/rpt_family="AluSx"

4195. 4245

/rpt_family="MER7A"

complement(4258. 4386)

/rpt_family="MLT2B3"

complement(4385. 4623)

/rpt_family="MLT2B3"

complement(4859. 4953)

/rpt_family="L1PA9"

complement(7027. 7223)

/rpt_family="L2"

7260. 7342

/rpt_family="(TATG)n"

complement(7349. 7541)

/rpt_family="L2"

complement(9158. 9165)

/note="<30 qual SNGL region"

complement(9182. 9491)

/rpt_family="AluY"

12672. 12990

/rpt_family="L1MB7"

13060. 13216

/rpt_family="L1MB7"

13402. 13694

/rpt_family="AluJo"

13750. 13778

/rpt_family="AT_rich"

complement(14039. 20161)

/rpt_family="L1PA5"

complement(21564. 21843)

/rpt_family="AluSp"

23292. 23499

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

[illegible]

Db	169575	ATCTGTTACGCGCTCTAGATGCGCTCTTTTAAATAATAATATTTATTTGTTGGCTT	169516
Qy	1441	TTGCTA-TTTTCTTTCTTATGCTCTTGAGAGGATATGCAAAATTTACTAGTGATCCA	1499
Db	169515	TTGCTATTTTCTTTCTTATGCTCTTGAGAGGATATGCAAAATTTACTAGTGATCCA	169456
Qy	1500	AAGAATAAACTTTGGCGTTGGCAATCTTTTCTCATCTATCTTTTGGCTTTATTTTATAA	1559
Db	169455	AAGAATAAACTTTGGCGTTGGCAATCTTTTCTCATCTATCTTTTGGCTTTATTTTATAA	169396
Qy	1560	TTCTGTTCTTGTTTTATAATGCTCTTTTATCTTTTGTGTTTACTTTGCTGTTCTTTT	1619
Db	169395	TTCTGTTCTTGTTTTATAATGCTCTTTTATCTTTTGTGTTTACTTTGCTGTTCTTTT	169336
Qy	1620	GTAATAATCCTCAGTAGAATGCTTAACCTTATTTGACATTCAGTCTTTCTCATTTCTACTAT	1679
Db	169335	GTAATAATCCTCAGTAGAATGCTTAACCTTATTTGACATTCAGTCTTTCTCATTTCTACTAT	169276
Qy	1680	GAGTATTTAGAGCCATAAAATTTCCCTTTTAACTTCCCTTTCCACTTCAACTACATCTCAC	1739
Db	169275	GAGTATTTAGAGCCATAAAATTTCCCTTTTAACTTCCCTTTCCACTTCAACTACATCTCAC	169216
Qy	1740	AAATTTGGATAGAGTAGTTTAAATTAATCATTTAGTATCTAAATATTTTAAATTTCTGTA	1799
Db	169215	AAATTTGGATAGAGTAGTTTAAATTAATCATTTAGTATCTAAATATTTTAAATTTCTGTA	169156
Qy	1800	TTTTCTTCTTTGATCCTGCAACTATTTTACAGTATTTTAAATCTGGAATAATAAGAT	1859
Db	169155	TTTTCTTCTTTGATCCTGCAACTATTTTACAGTATTTTAAATCTGGAATAATAAGAT	169096
Qy	1860	TGTTATTTGTTATTTGTTGATCTGATCTCTAAATTTGAATATTTAGATCAGATAATGTG	1919
Db	169095	TGTTATTTGTTATTTGTTGATCTGATCTCTAAATTTGAATATTTAGATCAGATAATGTG	169036
Qy	1920	GTTTGTAGGACATAATCCTTTGCAATTTGTTAGGCTTCTTTGGAACTTAATATGTC	1979
Db	169035	GTTTGTAGGACATAATCCTTTGCAATTTGTTAGGCTTCTTTGGAACTTAATATGTC	168976
Qy	1980	TCAATTTTATAGACGTTCTGTGTTCTTTGGGAAAAACATGATTTTGAATGTTGTTGG	2039
Db	168975	TCAGTTTATAGACGTTCTGTGTTCTTTGGGAAAAACATGATTTTGAATGTTGTTGG	168916
Qy	2040	TTTAATAATTTGTTATTTGATCATTGATTTGATTTGCTTATTTATTTGGCTGAATCTCCA	2099
Db	168915	TTTAATAATTTGTTATTTGATCATTGATTTGATTTGCTTATTTATTTGGCTGAATCTCCA	168856
Qy	2100	TTATCCTTAATGCTCTCTCAATTTTGTCTGCTTCTTTTATTAATTAGAGATAATGTTA	2159
Db	168855	TTATCCTTAATGCTCTCTCAATTTTGTCTGCTTCTTTTATTAATTAGAGATAATGTTA	168796
Qy	2160	AATTTATCTCACCTCAGTATAGTATGCTGTTTATATCTATATATAATAAATTTATAAT	2219
Db	168795	AATTTATCTCACCTCAGTATAGTATGCTGTTTATATCTATATATAATAAATTTATAAT	168736
Qy	2220	CCATAAATTTATGTTATGTAATTTGGAGACCTTATTTATCATATATAAACAAGATTTGTG	2279
Db	168735	CCATAAATTTATGTTATGTAATTTGGAGACCTTATTTATCATATATAAACAAGATTTGTG	168676
Qy	2280	ATGAATGACAGACTTATCTATTTAGTAGTACCTTTTATCTCGTCATATGTTATTTG	2339
Db	168675	ATGAATGACAGACTTATCTATTTAGTAGTACCTTTTATCTCGTCATATGTTATTTG	168616
Qy	2340	ACTTTGCTCAAAATTTTATTAATATTTGTTTGGTATTTCTTTTTCAGCGGGTTT	2399
Db	168615	ACTTTGCTCAAAATTTTATTAATATTTGTTTGGTATTTCTTTTTCAGCGGGTTT	168556
Qy	2400	ATGTCAGTCTGTTGCAATGGTGACACGCTGATTTTATTTAGACATGCTAGCTTTTAA	2459
Db	168555	ATGTCAGTCTGTTGCAATGGTGACACGCTGATTTTATTTAGACATGCTAGCTTTTAA	168496
Qy	2460	TTATTTCTTTTCCAAATTTCTTTTATTAATTTCTGATATACAAATTTTAGTCTACTTTT	2519

Db

168495

TTATTTCTTTTCCATTTTCAATTTTATAAATTTCTGATATACAATATTTAGTCACTTTT

168436

Qy

2520

ACCTTCTCTAGTGGAATTTT

2541

Db

168435

ACCTTCTCTAGTGGAATTTT

168414

RESULT 5

AP003400

135637 bp DNA linear PRI 15-MAR-2003

LOCUS

Homo sapiens genomic DNA, chromosome 11q, clone:RP11-643G5,

DEFINITION

complete sequence.

ACCESSION

AP003400

VERSION

AP003400.2

GI:15320508

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE

Homo sapiens genomic DNA

JOURNAL

Published Only in Database (2001)

REFERENCE

2 (bases 1 to 135637)

AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE

Direct Submission

JOURNAL

Submitted (08-MAR-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT

On Aug 27, 2001 this sequence version replaced gi:13366104.

FEATURES

Location/Qualifiers

1..135637

source

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11q"

/map="11q"

/clone="RP11-643G5"

ORIGIN

Query Match 93.2%; Score 2367; DB 9; Length 135637;

Best Local Similarity 96.8%; Pred. No. 0;

Matches 2470; Conservative 0; Mismatches 70; Indels 11; Gaps 5;

Qy

1

AGCATAATTTTCTTTTAACTTCTACCTAGTGTGATTCAGATGAGTTCAATTTCTCATGTGAAT

60

Db

132900

AGCATAATTTTCTTTTAACTTCTACCTAGTGTGATTCAGATGAGTTCAATTTCTCATGTGAAT

132959

Qy

61

CACAGAAAAAATATGGGAAATTTGGAATATGCGGATGTGGCAAGTCGTAGTTGATTT

120

Db

132960

CACAGAAAAAATATGGGAAATTTGGAATATGCGGATGTGGCAAGTCGTAGTTGATTT

133019

Qy

121

GTTTTCATTTTAGCTTCTATCCATGTGAGAAAGGAAATAACTACTCTAGTCACGTA

180

Db

133020

GTTTTCATTTTAGCTTCTATCCAGTGTGAGAAAGGAAATAACTACTCTAGTCACGTA

133079

Qy

181

CAATGTAATAATCTCCAGATCCCTCTCTCATCTCCAAATCCCTCCCAATTTGTAATCTG

240

Db

133080

CAATGTAATAATCTCCAGATCCCTCTCTCATCTCCAAATCCCTCCCAATTTGTAATCTG

133139

Qy

241

CACGTACCAAGACCTTATACATATAATGTTGAAAGAAAGTTTTCATAGCAGATATCCCTA

300

Db

133140

CTCTGGCCCAAGACCTTATACATATAATGTTGAAAGAAAGTTTTCATAGCAGATATCCCTA

133199

Qy

301

ACTTTCTTCTGGTCTTGAGGAAATACCTTTAAATTTCTTCTCACCAATTAATTTAGTAT

360

Db

133200

ACTTTCTTCTGGTCTTGAGGAAATACCTTTAAATTTCTTCTCACCAATTAATTTAGTAT

133259

Qy

361

TACTGTTTTCTTTTCTTCTGTTATCTTAAAGTACATCCCTTCCACATATACATAAATTTT

420

Db	133260	 TACTGTTTCTTTTCTGGTATCTTAAAGTACATCCTTTCCACAATATACTAAAAATTTT	133319
Qy	421	 TGCTGTATGATTTATATAACACATTTTTC--TTTAAAGATGAGCATGCAATTTTCTTT	479
Db	133320	 TGCTATATGATTTATATATAACACATTTTCTTTTGAGATGAGCAGGAGATTTTCTTT	133379
Qy	480	 CTTGATCTGAAGATGACATTCAGTGAATCTTAATGTTAGAATACTTTATTTTTTGA	539
Db	133380	 CTTGATCTGAAGATGACATTCAGTGAATCTTAATGTTAGAATACTTTATTTTTTGA	133439
Qy	540	 ACTTGTAAATATTAGCATTAGGTTTTAAATTCATATATTGATATATATTATGTTATATCT	599
Db	133440	 ACTTGTAAATATTAGCATTAGGTTTTAAATTCATATATTGATATATATTATGTTATATCT	133499
Qy	600	 CATAGCATGTTAGTCTGTGTTTGGCAATCTAATTTGTTGGGTATAAATTCAAATATCT	659
Db	133500	 CATAGCATGTTAGTCTGTGTTTGGCAATCTAATTTGTTGGGTATAAATTCAAATATCT	133559
Qy	660	 ACTTGCCCTCTAAAATGAGTGGTTTTTAATATTTTCTGAAGTAGGTTTTATATGCAATT	719
Db	133560	 ACTTGCCCTCTAAAATGAGTGGTTTTTAATATTTTCTGAAGTAGGTTTTATATGCAATT	133619
Qy	720	 AAATATTTTTTCTTTAACTTTCAAACCTCAAGGAAACACAGTTGGCCCTTGACCTGTT	779
Db	133620	 AAATATTTTTTCTTTAACTTTCAAACCTCAAGGAAACACAGTTGGCCCTTGACCTGTT	133679
Qy	780	 TGTGGAAAAATTTAAACTACTGCTTTAAATCTCTTTATGTTGTTGTAATAGCTATTTTAC	839
Db	133680	 TGTGGAAAAATTTAAACTACTGCTTTAAATCTCTTTATGTTGTTGTAATAGCTATTTTAT	133739
Qy	840	 GTCAATATAACAAATTTTATGTTTGTAAATGACTTTATTTGTTGTCATATGATAATTT	899
Db	133740	 GTCAATATAACAAATTTTATGTTTGTAAATGACTTTATTTGTTGTCATATGATAATTT	133799
Qy	900	 ATGTCATAGAACAAATTTTATGCTTGATATATGACTTTATTTGTTGTCATATGATAAC	959
Db	133800	 ATGTCATAGAACAAATTTTATGCTTGATATATGACTTTATTTGTTGTCATATGATAAC	133859
Qy	960	 TAGAATTTTTTGTGTTTGTACCGAGTCTTACTCTGTCCACCGAGCTGGAGTGAATGG	1019
Db	133860	 TAGAATTTTTTGTGTTTGTACCGAGTCTTACTCTGTCCACCGAGCTGGAGTGAATGG	133919
Qy	1020	 CATGGTCTCAGCTCAGTGCACACCTCCGCCCTCCGGGTTCAAGCCATTTCTCCACCTCAGC	1079
Db	133920	 CATGGTCTCAGCTCAGTGCACACCTCCGCCCTCCGGGTTCAAGCCATTTCTCCACCTCAGC	133979
Qy	1080	 CTCCCAAGTACCTGGACATCAGGCAATGACCAACCGCACCGGCTAAATTTTGTATTTT	1139
Db	133980	 CTCCCAAGTACCTGGACATCAGGCAATGACCAACCGCACCGGCTAAATTTTGTATTTT	134039
Qy	1140	 AGTAGAGACGTGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCCTGACCTCTGAATC	1199
Db	134040	 AGTAGAGACGTGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCCTGACCTCTGAATC	134099
Qy	1200	 CACCCGCTCGGCTCGCAAGTGTGGGATACAGGCGTAGCCATTTGTGCTCGCCGA	1259
Db	134100	 CACCCGCTCGGCTCGCAAGTGTGGGATACAGGCGTAGCCATTTGTGCTCGCCGA	134159
Qy	1260	 TTTTTTTTAAAAAGTATCTTATGTCAGTTTTTCAATAGTTTTTATTAAGTCAATTTTCC	1319
Db	134160	 TTTTTTTTAAAAAGTATCTTATGTCAGTTTTTCAATAGTTTTTATTAAGTCAATTTTCC	134219
Qy	1320	 ATTTGATGTAAAGCTTTCAAAATTTATAGTATGTTGTTCTAGTATTTTCTTATCTTTTGT	1379
Db	134220	 ATTTGATGTAAAGCTTTCAAAATTTATAGTATGTTGTTCTAGTATTTTCTTATCTTTTGA	134279
Qy	1380	 AATCTGTCAGCGCTCTGTAGATGTCCTCTTTTAAATAAATAATATTTATTTGTCGCT	1439
Db	134280	 AATCTGTCAGCATCTGCAGATGTCCTCTTTTGTATTAATTAATTTATTTATTTGTCCT	134339
Qy	1440	 TTTTGC--TATTTTTTTTCTATGCTCTTGTAGAGGATATGCAAAATTTACTAGTGATC	1497

RESULT 6	AP003122	137888 bp	DNA	linear	PRI 15-MAR-2003
LOCUS	AP003122	137888 bp	DNA	linear	PRI 15-MAR-2003
DEFINITION	Homo sapiens genomic DNA, chromosome 11q, clone:RP11-31312, complete sequence.				
ACCESSION	AP003122	GI:15320502			
VERSION	AP003122.2	GI:15320502			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.				
REFERENCE	Homo sapiens genomic DNA				
AUTHORS	Published Only in Database (2001)				
TITLE	2 (bases 1 to 137888)				
JOURNAL	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.				
REFERENCE	Direct Submission				
AUTHORS	Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)				
COMMENT	On Aug 27, 2001 this sequence version replaced gi:12597178.				
FEATURES	Location/Qualifiers				
source	1..137888				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="11"				
	/map="11q"				
	/clone="RP11-31312"				
ORIGIN					
	Query Match	93.2%;	Score 2367;	DB 9;	Length 137888;
	Best Local Similarity	96.8%;	Pred. No. 0;		
	Matches 2470;	Conservative	0;	Mismatches	70;
				Indels	11;
				Gaps	5;
Qy	1	AGCATAATTTTCTTTAACTTACCTAGTAGTGAATCAGATGAGTTCAATTCCTCATGTGAAT	60		
Db	104	AGCATAATTTTCTTTAACTTACCTAGTAGTGAATCAGATGAGTTCAATTCCTCATGTGAAT	163		
Qy	61	CACAGAAAATATGGAAATTTGGAAATATGGGATCTGGCAATCTGGCAATCTGGATTTT	120		
Db	164	CACAGAAAATATGGAAATTTGGAAATATGGGATCTGGCAATCTGGCAATCTGGATTTT	223		
Qy	121	GGTTTCATTTTGTAGCTTCTATCCATGTCTAGAAAGGAAATAACTACTGTCTAGTCAAGTA	180		
Db	224	GGTTTCATTTTGTAGCTTCTATCCAGTGTCTAGAAAGGAAATAACTACTGTCTAGTCAAGTA	283		
Qy	181	CAATGTAAATCTGCCAGATCCCTCTCATCTCCAAATCCCTCCCATTTGTATTAATCTG	240		
Db	284	CAATGTAAATCTGCCAGATCCCTCTCATCTCCAAATCCCTCCCATTTGTATTAATCTG	343		
Qy	241	CACATGACCAAGACCTATACATAATATGTTGAAAGAAAGTTTGTATAGCAGATATCCCTTA	300		
Db	344	CTCTGGCCAGACCTATACATAATATGTTGAAAGAAAGTTTGTATAGCAGATATCCCTTA	403		
Qy	301	ACTTCTCTGGCTCTGAAGAAATACCTTTAATCTTCTTCCACCATTAATATGATTT	360		
Db	404	ACTTCTCTGGCTCTGAAGAAATACCTTTAATCTTCTTCCACCATTAATATGATTT	463		
Qy	361	TACTGTTTTCTTTTCTGTTATCTTAAAGTACATCTTTCCACATATACTAAATTTT	420		
Db	464	TACTGTTTTCTTTTCTGTTATCTTAAAGTACATCTTTCCACATATACTAAATTTT	523		
Qy	421	TGCTGTATGGAATTTATATTAACACATTTTTC--TTTAAGATGAGCATGGCATTTTCTTT	479		


```
* 31650 34631: contig of 2982 bp in length
* 34632 34731: gap of unknown length
* 34732 38468: contig of 3737 bp in length
* 38469 38568: gap of unknown length
* 38569 42337: contig of 3769 bp in length
* 42338 42438: gap of unknown length
* 42439 47543: contig of 5106 bp in length
* 47544 47643: gap of unknown length
* 47644 56048: contig of 8405 bp in length
* 56049 56148: gap of unknown length
* 56149 64671: contig of 8523 bp in length
* 64672 67711: gap of unknown length
* 67712 73100: contig of 8329 bp in length
* 73101 73201: gap of unknown length
* 73202 81814: contig of 8614 bp in length
* 81815 81914: gap of unknown length
* 81915 91133: contig of 9219 bp in length
* 91134 91233: gap of unknown length
* 91234 101427: contig of 10194 bp in length
* 101428 101528: gap of unknown length
* 101529 113041: contig of 11514 bp in length
* 113042 113141: gap of unknown length
* 113142 125504: contig of 12363 bp in length
* 125505 125604: gap of unknown length
* 125605 140510: contig of 14906 bp in length
* 140511 140611: gap of unknown length
* 140612 158204: contig of 17594 bp in length
* 158205 158304: gap of unknown length
* 158305 192648: contig of 34344 bp in length.
Location/Qualifiers
1. .192648
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="11"
/clone="RP11-31312"
1. .1413
/note="assembly_name:Contig10"
1514. .2978
/note="assembly_name:Contig11"
3079. .4859
/note="assembly_name:Contig12"
4960. .6723
/note="assembly_name:Contig13"
6824. .9068
/note="assembly_name:Contig14"
9169. .11510
/note="assembly_name:Contig15"
11611. .15431
/note="assembly_name:Contig16"
15332. .18469
/note="assembly_name:Contig17"
18570. .22092
/note="assembly_name:Contig18"
22193. .25250
/note="assembly_name:Contig19"
25351. .28352
/note="assembly_name:Contig20"
28453. .31549
/note="assembly_name:Contig21"
31650. .34631
/note="assembly_name:Contig22"
34732. .38468
/note="assembly_name:Contig23"
38569. .42337
/note="assembly_name:Contig24"
42438. .47543
/note="assembly_name:Contig25"
47644. .56048
/note="assembly_name:Contig26"
56149. .64671
/note="assembly_name:Contig27"
64772. .73100
```

```
misc_feature /note="assembly_name:Contig28" 73201. .81814
misc_feature /note="assembly_name:Contig29" 81915. .91133
misc_feature /note="assembly_name:Contig30" 91234. .101427
misc_feature /note="assembly_name:Contig31 clone_end:86 vector_side:left" 101528. .113041
misc_feature /note="assembly_name:Contig32" 113142. .125504
misc_feature /note="assembly_name:Contig33" 125605. .140510
misc_feature /note="assembly_name:Contig34" 140611. .158204
misc_feature /note="assembly_name:Contig35" 158305. .192648
misc_feature /note="assembly_name:Contig36"

ORIGIN
Query Match 93.2%; Score 2367; DB 2; Length 192648;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 2470; Conservative 0; Mismatches 70; Indels 11; Gaps 5;

QY 1 AGCATAAATTTTCTTTTAATCTTACCTAGTATTCAGATGAGTTCATTTCTCATGTGAAT 60
Db 174965 AGCATAAATTTTCTTTTAATCTTACCTAGTATTCAGATGAGTTCATTTCTCATGTGAAT 174906

QY 61 CACAGAAAAAATATGCGAAATTTGGAATATGTGGGATGTGGCAAGTCGTAGTTCATTT 120
Db 174905 CACAGAAAAAATATGCGAAATTTGGAATATGTGGGATGTGGCAAGTCGTAGTTCATTT 174846

QY 121 GGTTCATTTTAGCTTCTATCCATGTCAGAAAAGGAAATAACTACTGTCTAGTCAGTGA 180
Db 174845 GGTTCATTTTAGCTTCTATCCATGTCAGAAAAGGAAATAACTACTGTCTAGTCAGTGA 174785

QY 181 CAATGTAATCTGCCAGATCCCTCCATCTCCAAATCCCTCCATCTCCATCTGTTACTG 240
Db 174785 CAATGTAATCTGCCAGATCCCTCCATCTCCAAATCCCTCCATCTGTTACTGTTACTG 174726

QY 241 CACTGACCAAGACCTATACATATATATGTTGAAAAGAGTTTGTAGACAGATATCCTTA 300
Db 174725 CACTGACCAAGACCTATACATATATATGTTGAAAAGAGTTTGTAGACAGATATCCTTA 174666

QY 301 ACTTCTCTCGTCTTTGAAGGAAATACCTTTAATCTTTCTCAACATTAATTATGTAATT 360
Db 174665 ACTTCTCTCGTCTTTGAAGGAAATACCTTTAATCTTTCTCAACATTAATTATGTAATT 174606

QY 361 TACTGTTTTCTTTTCTGTTATCTTAAAGTACATCTTCCACATATACATAAATTTT 420
Db 174605 TACTGTTTTCTTTTCTGTTATCTTAAAGTACATCTTCCACATATACATAAATTTT 174546

QY 421 TGCTGTATGATTTATTTAAACACATTTTTC-TTTAAGATGAGCATGCGATTTTCTTT 479
Db 174545 TGCTGTATGATTTATTTAAACACATTTTTC-TTTAAGATGAGCATGCGATTTTCTTT 174486

QY 480 CTTGATCTGAAGATGACATTCAGTGATATCTAATGTTAGAAATCTTATATTTTGA 539
Db 174485 CTTGATCTGAAGATGACATTCAGTGATATCTAATGTTAGAAATCTTATATTTTGA 174426

QY 540 ACTTGTAAATATAGCATTAGGGTTTAAATTCATATATATATATATATATATTTTACT 599
Db 174425 ACTTGTAAATATAGCATTAGGGTTTAAATTCATATATATATATATATATATTTTACT 174366

QY 600 CATAGCATGTTTAGTCTGTGTTTTTTCGAATCTAATTTTGGTGTATAAATTCATATCT 659
Db 174365 CATAGCATGTTTAGTCTGTGTTTTTTCGAATCTAATTTTGGTGTATAAATTCATATCT 174306

QY 660 ACTTGCTCTCTAAAATGAGTTGGGTTTTTAATATTTTCTGAAGTAGGTTTATGCAATT 719
Db 174305 ACTTGCTCTCTAAAATGAGTTGGGTTTTTAATATTTTCTGAAGTAGGTTTATGCAATT 174246
```

QY	720	AAATTATTTTTTCCCTTTAAACCTTTCAAACCTCAAGGAAAAACCAAGTTGGCCCTTGACTCTGTT	779
Db	174245	AAATTATTTTTTCCCTTTAAACCTTTCAAACCTCAAGGAAAAACCAAGTTGGCCCTTGACTCTGTT	174186
QY	780	TGTGAAAAATTTTAAACCTACTGGTTTAAATTTCTTTATTCGGTGTGAATATGACTATTTTAC	839
Db	174185	TGTGAAAAATTTTAAACCTACTGGTTTAAATTTCTTTATTCGGTGTGAATATGACTATTTTAT	174126
QY	840	GTCAATATAACAATTTTTTATTTGTTTGTAAATGACCTTTATTTGTTTGTTCATATGATAATTTT	899
Db	174125	GTCAATATAACAATTTTTTATTTGTTTGTAAAGACCTTTATTTGTTTGTTCATATGATAATTTT	174066
QY	900	ATGTCATAGAACAAATTTTTTATTTGCTGTGATATGACTTTTATTTGTTATATGCGCTATACAAC	959
Db	174065	ATGTCATAGAACAAATTTTTTATTTGTTTGTATATGACTTTTATTTGTTATATGCGCTATACAAC	174006
QY	960	TAGATTTTTTTTTTGTTTTTTGACCCAGACTTACTCTGTCAACCCAGCGCTGGAGTGAATGG	1019
Db	174005	TAGATTTTTTTTTTGTTTTTTGACCAAGCTTACTCTGTCAACCCAGCGCTGGAGTGAATGG	173946
QY	1020	CATGGTCTCAGCTCACTGCAACCTCGCGCTCCCGGGTTCAAGCCATTCCTCCACCTCAGC	1079
Db	173945	CATGGTCTCAGCTCACTGCAACCTCGCGCTCCAGGTTCAAGCAATTCCTCCGCTCAGC	173886
QY	1080	CTCCCAAGTAGCTGGGACTACAGGCAATGAGCCACCGCACCCGGCTAAATTTTTGTATTTTT	1139
Db	173885	CTCCCAAGTAGCTGGGACTACAGGCAAGCTGCCACCGCACCCGGCTAAATTTTTGTATTTTT	173826
QY	1140	AGTAGAGACGTGGTTCCACTATGTTGGCCAGCGCTGATCTCGAACTCCTGACCTTGTAATC	1199
Db	173825	AGTAGAGACCTGGTTTCCACTATGTTGGCCAGCGCTGGTCTCGAACTCCTGAACTCGTAATC	173766
QY	1200	CACCGCGCTCGCGCTGCCCAAGTGTGGGAATACAGGGGTGAGCCATTGTGCTCGCCGA	1259
Db	173765	CACCGCGCTCGCGCTGCCCAAGTGTGGGAATACAGGGGTGAGCCATTGTGCTCGCCGA	173706
QY	1260	TTTTTTAAAAAATGTAATCTTATGTGCTGTTTTCAATGATTTTATTTAAAAATGCAATTTTCC	1319
Db	173705	TTTTTTAAAAAATGTAATCTTATGTGCTGTTTTCAATGATTTTATTTAAAAATGCAATTTTCC	173646
QY	1320	ATTGTAGCTAAGCTTCCAAATTTATAGTATAGTTGTTCCCTAGTATTTTCTTAATCTTTTGT	1379
Db	173645	ATTTCATGTAAGCTTCCAAATTTTATAGTATAGTTGTTCCCTAGTATTTTCTTATCTTTTAA	173586
QY	1380	AATCTGTTCAAGCGTCTGAGATGTGCCCTCTTTTAAATAAATAATTAATTTGTTTGGCT	1439
Db	173585	AATCTGTTCAAGCACTGCGAGATGTGCCCTCTTTTGATTAATTAATTAATTTTGTGCT	173526
QY	1440	TTTGG--TATTTTTTTCTTAATGCTCTGTGAGAGGATATGTCAAATTTACTAGTGTATC	1497
Db	173525	TTTGTCTATATTTTTTTCTTAATGCTCTGTGAGAGGATATGTCAAATTTACTAGTGTATC	173466
QY	1498	CAAGAATAAATTTTGGCGTTGGCAATCTTTTCTCATCTATCTTTGCTTTATTTATTTAT	1557
Db	173465	CAAGAATAAATTTTGGCTTTGGCAATCTTTTCTCATCTATCTTTGCTTTATTTATTTAT	173406
QY	1558	AAATCTGTTCTTGTTTTTATAATTGCCCTCTTTTATCTCTTTTGTTGTTACTTTTGCTGTTCT	1617
Db	173405	AAATCTGTTCTTGTTTTTATAATTGCCCTCTTTTATCTCTTTTGTTGTTACTTTTGCTGTTCT	173346
QY	1618	TTGTAAATCTCTAGTGAAGTCTTAATTCATATGCAATTCAGTCTTTCTTCATTTCTACT	1677
Db	173345	TTGTAAATCTCTAGTGAAGTCTTAATTCATATGCAATTCAGTCTTTCTTCATTTCTACT	173286
QY	1678	ATGAGTATTTAGAGCCATTAATTTCCCTTTTAACTTCCCTTTCACCTCAACTCACTCTC	1737
Db	173285	ATGAGCATTTAGAGCCATAATTTCCCTTTTAACTTCCCTTTCACCTCAACTCACTCTC	173226
QY	1738	ACAAATTTGGATTAGGAGTAGTTTAAATTAATCATTAGTATCTAAATATTTTTTAATTTCTG	1797
Db	173225	ACAAATTTGGATTAGGAGTAGTTTAAATTAATCATTAGTATGAATAATTTTTTAATTTCTG	173166
QY	1798	TATTTTCTCTTTGATCCTCGAACTATTTTCAAGTATTTTTTTAAAAATCCTGAAATATAAG	1857

Db	173165	TATTTCTCTTTGATCTCGAACTGTTTACAG-ATTTTTTAAATCCTGAAATGAAAG	173107
Qy	1858	ATTGTTATGTTTATTTGTTGATCTGATCTCTAAATTTGAATATATTGAGATCAGATAATG	1917
Db	173106	ATTGTTATTTATTTATTTGTTGATCTGATCTCTAAATTTGAACATATTGAGATCAGATAATG	173047
Qy	1918	TGGTTTGTAGGACACTAATCCTTTTGACAAATGTTGAGGGCTTCCTTTGGAACTTAATATGT	1977
Db	173046	TGGTTTGTAGGACACTAATCCTTTTGACAAATGTTGAGGGCTTCCTTTGGAACTTAATATGT	172989
Qy	1978	GCTCAATTTTTATACACGTTCTGCTTCTTTTGGGAAAAACAATGATTTATGATGGTTGTTT	2037
Db	172986	GCTCAATTTTTATATAATGTTCTGCTTCTTTTGGGAAAAACAATGATTTATGATGGTTGTTT	172927
Qy	2038	GGTTTAAATTTTGTATTTGATTTGACATTTAGTTGAGTTTGTCTTATTTTGGCTGAAATCTC	2097
Db	172926	GGTTTAAATTTTATTTATTTGATTTGACATTTAGTTGAGTTTGTCTTATTTTGGCTGAAATCTC	172867
Qy	2098	CATTATCCCTTAATGCTCTCTCATTTTGTCTGCTTCCCTTTTAAATTTAGAGATAAATGT	2157
Db	172866	CATTATCCCTTAATGCTCTCTCATTTTGTCTGCTTCCCTTTTAAATTTAGAGATAAATGT	172807
Qy	2158	TAAATATCTCACCTCAGTATGATGCTGTTTATATATATATATAATAAATTTATAA	2217
Db	172806	TAAATATCTCACCTCAGTATGATGCTGTTTATATATATATATAATAAATTTATAA	172747
Qy	2218	TTCATATAATTTATGTTATGATATAATTTGGAGACCTATTATCATATATAAAC-----AG	2271
Db	172746	TTCATATAATTTATGTTATGATATAATTTGGAGACCTATTATCATATATAAAC-----AG	172687
Qy	2272	AATTTGTTGATCAAGATGACAGACTTATCTTATGATGATAGTACCTTTTTTATCTCGTCATAAT	2331
Db	172686	AATTTGTTGATCAAGATGACAGACTTATCTTATGATGATAGTACCTTTTTTATCTCGTCATAAT	172627
Qy	2332	GTTATTTGACCTTCTCTTAAATTTTTTTTT-AATTAATTTTGTGTTATTTCTTTTTC	2390
Db	172626	GTTATTTGACCTTCTCTTAAATTTTTTTTT-AATTAATTTTGTGTTATTTCTTTTTC	172567
Qy	2391	ACGGGTTTATGTCAGTCTTGTCAATTTGGTACAGCTGATTTTATTAGACATGCTAC	2450
Db	172566	AGTAGGTTTATGTCAGTCTTGTCAATTTGGTACAGCTGATTTTATTAGACATGCTAC	172507
Qy	2451	GCCTTTTAAATTTATCTTTTTCGATTTTCATTTTATTAATTCGATATACAAATTTAG	2510
Db	172506	GCCTTTTAAATTTATCTTTTTCGATTTTCATTTTATTAATTCGATATACAAATTTAG	172447
Qy	2511	GTCACCTTTTACCTTCTCTAGTGTGAATTTT	2541
Db	172446	GTCACCTTTTACCTTCTCTAGTGTGAATTTT	172416
RESULT 8			
AC023784			
LOCUS		117521 bp	DNA
DEFINITION		Homo sapiens chromosome 11 clone CTD-2026G22, WORKING DRAFT	linear HTG 07-JUL-2000
ACCESSION		AC023784	
VERSION		AC023784.3	GI:7230809
KEYWORDS		HTG; HTGS PHASE1; HTGS_DRAFT.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Theria; Primates; Catarrhini; Hominiidae; Homo.	
TITLE		Waterston, R.H.	
JOURNAL		The sequence of Homo sapiens clone	
REFERENCE		Unpublished	
AUTHORS		2 (bases 1 to 117521)	
TITLE		Waterston, R.H.	
JOURNAL		Direct Submission	
		Submitted (17-FEB-2000) Genome Sequencing Center, Washington	
		University School of Medicine, 4444 Forest Park Drive, SE, Atlanta, GA 30322	

MO 63108, USA
On Mar 13, 2000 this sequence version replaced gi:7021753.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H.MS2026G22
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 108941 bases at least Q40
Consensus quality: 111448 bases at least Q30
Insert size: 132000; agarose-fp
Insert size: 115321; sum-of-contigs
Quality coverage: 3.18 in Q20 bases; agarose-fp
Quality coverage: 3.67 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1436: contig of 1436 bp in length
* 1437 1536: gap of unknown length
* 1537 2618: contig of 1082 bp in length
* 2619 2718: gap of unknown length
* 2719 4121: contig of 1403 bp in length
* 4122 4221: gap of unknown length
* 4222 5276: contig of 1055 bp in length
* 5277 5376: gap of unknown length
* 5377 6917: contig of 1541 bp in length
* 6918 7017: gap of unknown length
* 7018 8794: contig of 1777 bp in length
* 8795 8894: gap of unknown length
* 8895 13539: contig of 4645 bp in length
* 13540 13639: gap of unknown length
* 13640 17094: contig of 3455 bp in length
* 17095 17194: gap of unknown length
* 17195 20927: contig of 3733 bp in length
* 20928 21027: gap of unknown length
* 21028 24808: contig of 3781 bp in length
* 24809 24908: gap of unknown length
* 24909 28904: contig of 3996 bp in length
* 28905 29004: gap of unknown length
* 29005 32266: contig of 3262 bp in length
* 32267 32366: gap of unknown length
* 32367 36987: contig of 4621 bp in length
* 36988 37087: gap of unknown length
* 37088 40336: contig of 3249 bp in length
* 40337 40436: gap of unknown length
* 40437 46367: contig of 5931 bp in length
* 46368 46467: gap of unknown length
* 46468 52979: contig of 6512 bp in length
* 52980 53079: gap of unknown length
* 53080 58971: contig of 5892 bp in length
* 58972 59071: gap of unknown length
* 59072 65824: contig of 6753 bp in length
* 65825 65924: gap of unknown length
* 65925 72767: contig of 6843 bp in length
* 72768 72867: gap of unknown length
* 72868 80928: contig of 8061 bp in length
* 80929 81028: gap of unknown length
* 81029 92294: contig of 11266 bp in length
* 92295 92394: gap of unknown length

* 92395 103346: contig of 10952 bp in length
* 103347 103446: gap of unknown length
* 103447 117521: contig of 14075 bp in length.
FEATURES
Location/Qualifiers
1..117521
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/clone="CTD-2026G22"

ORIGIN

Query Match 84.5%; Score 2148; DB 2; Length 117521;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 2179; Conservative 0; Mismatches 115; Indels 2; Gaps 2;
QY 247 CCAAGACCTATACAATAATCTTGAAGAAGAGTTTTCATAGCAGATATCTTAACTTTC 306
Db |||||
QY 24788 CTAAGTCGTACTAATCAAGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 24847
Db |||||
QY 307 TTCTGGTCTTGAAGGAATACCTTTAAATCTTCTCACCAATTAATATGATGTAATCTGT 366
Db NNN 24907
QY 367 TTCTTTTCTCGTTTCTTAAAGTACATCCCTTCCACATATACATAAAATTTTTCCTGT 426
Db |||||
QY 24908 NTTCCTTTTCTGGTTATCTTAAAGTACATCCCTTCCACATATACATAAAATTTTTCCTGT 24967
Db |||||
QY 427 ATGGATTATATTAACAACATTTTCTTTAAGATGAGCATGCGATTTTCTTCTTGTATC 486
Db |||||
QY 24968 ATGGAATATATTAACAACATTTTCTTTAAGATGAGCATGCGATTTTCTTCTTGTATC 25027
Db |||||
QY 487 TGAAGATGATTCAGTGTATATCTTAATGTTTGAATATCTTATATTTTGAATTTGTA 546
Db |||||
QY 25028 TGAAGATGATTCAGTGTATCTTAAATGTTTGAATATCTTATATTTTGAATTTGTA 25086
Db |||||
QY 547 ATATTAGCATTTAGGTTTAAATTCATATATTTGATATATATTTGATATATTTGATAT 606
Db |||||
QY 25087 ATATTAGCATTTAGGTTTAAATTCATATATTTGATATATATTTGATATATTTGATAT 25146
Db |||||
QY 607 TGTTAGTCTGTCTTTTTCGAATCTATTTGTTGGTTTAAATTTCAATTTCTTCTTGTCC 666
Db |||||
QY 25147 TGTTAGTCTGTCTTTTTCGAATCTATTTGTTGGTTTAAATTTCAATTTCTTCTTGTCC 25206
Db |||||
QY 667 TTCTAAATAGTGTGGTTTAAATTTCTGAAGTAGTTTATTTGAAATTTAAATTTAT 726
Db |||||
QY 25207 TTCTAAATAGTGTGGTTTAAATTTCTGAAGTAGTTTATTTGCAATTTAAATTTAT 25266
Db |||||
QY 727 TTTTTCCTTTAACTTTCAAACTCAAGGAACACAGTTCGCTTGTCTGTTTGTGGAA 786
Db |||||
QY 25267 TTTTTCCTTTAACTTTCAAACTCAAGGAACACAGTTCGCTTGTCTGTTTGTGGAA 25326
Db |||||
QY 787 AATTTTAACTACTGTGTTTAAATTTCTTATTTGTTGTTAATATGACTATTTTACGTCA 846
Db |||||
QY 25327 AATTTTAACTACTGTGTTTAAATTTCTTATTTGTTGTTAATATGACTATTTTACGTCA 25386
Db |||||
QY 847 AACAAATTTATTTGTTTAAATGACTTTTATTTGTTTCAATATGATATTTTATGTCAT 906
Db |||||
QY 25387 AACAAATTTATTTGTTTAAATGACTTTTATTTGTTTCAATATGATATTTTATGTCAT 25446
Db |||||
QY 907 AGAACAATTTTATTTGTTTAAATGACTTTTATTTGTTTCAATATGATATTTTACGTCA 966
Db |||||
QY 25447 AGAACAATTTTATTTGTTTAAATGACTTTTATTTGTTTCAATATGATATTTTACGTCA 25506
Db |||||
QY 967 TTTTGTGTTTGTGACCGAGTCTTACTCTGTCAACCGAGGTGAGGTGATGATGATGATG 1026
Db |||||
QY 25507 TTTTGTGTTTGTGACCGAGTCTTACTCTGTCAACCGAGGTGAGGTGATGATGATGATG 25566
Db |||||
QY 1027 TCAGCTCACTGCAACCTCCGCTCCGCGGTTCAGGCAATTTCTCCACCTCAGCCTCCCAA 1086
Db |||||
QY 25567 TCAGCTCACTGCAACCTCCGCTCCGCGGTTCAGGCAATTTCTCCACCTCAGCCTCCCAA 25626
Db |||||
QY 1087 GTAGCTGGGACTACAGGCATGAGCCACCGCCGCTAATTTTGTATTTTAGTAGAG 1146

[illegible]

Foundation, Box ND50, 9500 Euclid Ave., Cleveland, OH 44195, USA

FEATURES
Location/Qualifiers
source

1. 1669
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q14.3"
misc_feature
1. 1669
/note="prostate-specific membrane antigen-like enhancer
region; PSMA-like; FOUHI-like"

ORIGIN

Query Match 57.5%; Score 1461.4; DB 9; Length 1669;
Best Local Similarity 95.3%; Pred. No. 2.5e-215;
Matches 1596; Conservative 0; Mismatches 61; Indels 18; Gaps 8;

QY 664 GCCTCTAAATAGAGTGGGTTTAAATATTTCTGAAGTAGGTTTATGCAATTAAT 723
Db 1 GCCTCTAAATAGAGTGGGTTTAAATATTTCTGAAGTAGGTTTATGCAATTAAT 60
QY 724 TATTTTCTTAACTTCAACCTCAAGGAAACAGTGGCTTGAATCTGTTTGG 783
Db 61 TATTTTCTTAACTTCAACCTCAAGGAAACAGTGGCTTGAATCTGTTTGG 120
QY 784 GAAATTTTAACTACTGGTTTAAATTTCTTTATTTGTTGAATGACTATTTACGTCA 843
Db 121 GAAATTTTAACTACTGGTTTAAATTTCTTTATTTGTTGAATGACTATTTATGTC 180
QY 844 TATAACAAATTTTATTTGTTTAAAGACTTTATTTGTTGTCATATGATATTTATGT 903
Db 181 TATAACAAATTTTATTTGTTTAAAGACTTTATTTGTTGTCATATGATATTTATGT 240
QY 904 CATAGACAAATTTTATTTGTTTCAATATGACTTTATTTGTTATGCTATACAACTAGA 963
Db 241 CATAGACAAATTTTATTTGTTTCAATATGACTTTATTTGTTATGCTATACAACTAGA 300
QY 964 TTTTATTTGTTTGTGACGAGTCTTACTCTCTCACCCAGGCTGGAGTGAATGCGCATG 1023
Db 301 TTTTATTTGTTTGTGACCAAGCTTACTCTCTCACCCAGGCTGGAGTGAATGCGCATG 360
QY 1024 GTCTAGCTCACTGCAACCTCGCTCCCGGTTCAAGCCATCTTCCACTCAGCTCC 1083
Db 361 GTCTAGCTCACTGCAACCTCGCTCCCGGTTCAAGCCATCTTCCCGCTCAGCTCC 420
QY 1084 CAAGTACCTGGGACTACAGGACGAGCCACCGCCAGCGCTAAATTTTGTATTTTAGTA 1143
Db 421 CAAGTACCTGGGACTACAGGACGAGCCACCGCCAGCGCTAAATTTTGTATTTTAGTA 480
QY 1144 GAGACGTGGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCCTGACCTTGTATCCACC 1203
Db 481 GAGACGTGGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCCTGAACTCGTAAATCCACC 540
QY 1204 CGCTCGGCTGCAAGAGTGGGATTAAGGGTGGAGGATGAGGCTGGCGGATTT 1263
Db 541 CGCTCGGCTGCAAGAGTGGGATTAAGGGTGGAGGATGAGGCTGGCGGATTT 600
QY 1264 TTAATAAATCTATCTTATGTCAGTTTCTTCAAGTTTATTTAAATGCAATTTCCATTT 1323
Db 601 TTAATAAATCTATCTTATGTCAGTTTCTTCAAGTTTATTTAAATGCAATTTCCATTT 660
QY 1324 GATTAAGCTTTCAAAATTTATAGTATAGTTGTTCTTAGTATTTTCTTATCTTTGTAATC 1383
Db 661 CAATTAAGCTTTCAAAATTTATAGTATAGTTGTTCTTAGTATTTTCTTATCTTTTAAATC 719
QY 1384 TGTTACAGCTGTAGATGCGCTTTTAAATTAATTAATATTTATTTGTTGCGCTTTG 1443
Db 720 TGTTACAGCTGTAGATGCGCTTTTAAATTAATTAATATTTATTTGTTGCGCTTTG 778
QY 1444 CTATTTTCTTCTTATGCTTTCAGAGGATATGTCAAATTTACTAGTGTATCAAGA 1503
Db 779 CTATTTTCTTCTTATGCTTTCAGAGGATATGTCAAATTTACTAGTGTATCAAGA 838

QY 1504 ATAACTTTGGGTTGGCAATCTTTCTCACTATCTTTGCTTTATATTTTAAATCT 1563
Db 839 ATAACTTTGGGTTGGCAATCTTTCTCACTATCTTTGCTTTATATTTTAAATCT 898
QY 1564 GTTCTTTGTTTATAATTTGCTCTTTTATCTTTCTTTGTTGTTTACTTTGCTTTCTTTGTA 1623
Db 899 GTTCTTTGTTTATAATTTGCTCTTTTATCTTTCTTTGTTGTTTACTTTGCTTTCTTTGTA 958
QY 1624 AATCTCTAGTAGAATGCTTAACTTATGACATTCAGTCTTTCTTCAATTTCTACTAGT 1683
Db 959 AATCTCTAGTAGAATGCTTAACTTATGACATTCAGTCTTTCTTCAATTTCTACTAGT 1018
QY 1684 ATTTAGAGCCATAAATTTCCCTTTAACTTCCCTTTCCACTTCAATCTACATCTCAAAAT 1743
Db 1019 ATTTAGAGCCATAAATTTCCCTTTAACTTCCCTTTCCACTTCAATCTACATCTCAAAAT 1078
QY 1744 TTGGATTTAGGAGTAGTTTAAATATCATTTAGTATCTAAATATTTTAAATTTCTGTATTT 1803
Db 1079 TTGGATTTAGGAGTAGTTTAAATATCATTTAGTATGTAATATTTTAAATTTCTGTATTT 1138
QY 1804 CTCTTTTGATCTGCAACTTATTTACAGTATTTTAAATCTTAAATCTGTAATTAAGATGTT 1863
Db 1139 CTCTTTTGATCTGCAACTTATTTACAG-ATTTTAAATCTGTAATCTGTAATCTGTAAT 1196
QY 1864 ATTTGTTATTTGTTGATCTGATCTTAAATTTGAATATTTTGAATCAGATTAATGCTGTT 1923
Db 1197 ATTTGTTATTTGTTGATCTGATCTTAAATTTGAATTTGATGATGATGATGATGATGTT 1256
QY 1924 GTAGGACACTAAATCTCTTTGACAAATTTGAGGCTTCTTTGGAACCTAATATGCTCAA 1983
Db 1257 GTAGGACACTAAATCTCTTTGACAAATTTGAGGCTTCTTTGGAACCTAATATGCTCAA 1316
QY 1984 TTTTATAGACGTTCTGTTTCTTTTGGGAAACAAATGATTTTGAATGTTGTTGTTTGA 2043
Db 1317 -TTTTATAATGTTCTGTTTCTTTTGGGAAACAAATGATTTTGAATGTTGTTGTTTGA 1375
QY 2044 ATATTTTGTATTTGTACATTTAGTTGAGTTTGTCTTATTTTGGCTGAAATCTCCATTT 2103
Db 1376 ATATTTTGTATTTGTACATTTAGTTGAGTTTGTCTTATTTTGGCTGAAATCTCCATTT 1435
QY 2104 CCTTAATGCTCTCTCTCTTTGCTTTGCTTTTAAATTTAGAGATAAATGTTAAAT 2163
Db 1436 CCTTAATGCTCTCTCTCTTTGCTTTGCTTTTAAATTTAGAGATAAATGTTAAAT 1495
QY 2164 ATCTACCTCACTATAGTGAATGCTTTTATCTATATATATAAATTTATATTTCCAT 2223
Db 1496 A-CTCACCTCACTATAGTGAATGCTTTTATCTATATATATAAATTTATATTTCCAT 1554
QY 2224 AAATTTATGTTATGATTAATTTGGAGACCTTATTTATCATATATAAAC-----AGAATGT 2277
Db 1555 AAATTTATGTTATGATTAATTTGGAGACCTTATTTAAATATACAAATTAAGATAGT 1614
QY 2278 TGATGAAATGACAG-----ACTTATCTTATGTTAGTGGCTTTTATCTGTC 2326
Db 1615 TGATGAAATGACAGCTTATCTTATCTTATGTTAGTGGCTTTTATCTGTC 1669

RESULT 11

AC074003

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AC074003 246865 bp DNA linear HTG 24-SEP-2000
Homo sapiens chromosome 2 clone RP11-163G14, *** SEQUENCING IN
PROGRESS ***, 53 unordered pieces.

AC074003

AC074003.3 GI:10281455

HTG; HTGS PHASE1.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Waterston,R.H.

The sequence of Homo sapiens clone

Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 246865)

Waterston,R.H.
Direct Submission
Submitted (09-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

On Sep 24, 2000 this sequence version replaced gi:9186850.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

----- Project Information -----

Center project name: H.NH0163G14

----- Summary Statistics -----

Sequencing vector: M13; 100%

Sequencing vector: plasmid; 0%

Chemistry: Dye-terminator; 100% of reads

Chemistry: Dye-terminator Big Dye; 0% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 213368 bases at least Q40

Consensus quality: 223527 bases at least Q30

Consensus quality: 229418 bases at least Q20

Insert size: 170000; agarose-fp

Insert size: 241665; sum-of-contigs

Quality coverage: 4.83 in Q20 bases; agarose-fp

Quality coverage: 3.37 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 12900: contig of 12900 bp in length
12901 13000: gap of unknown length
13001 23006: contig of 10006 bp in length
23007 23106: gap of unknown length
23107 35032: contig of 11926 bp in length
35033 35132: gap of unknown length
35133 50532: contig of 15400 bp in length
50533 50632: gap of unknown length
50633 64756: contig of 14124 bp in length
64757 64856: gap of unknown length
64857 80978: contig of 16122 bp in length
80979 101159: contig of 20081 bp in length
101160 101239: gap of unknown length
101240 102865: contig of 1606 bp in length
102866 102965: gap of unknown length
102966 104593: contig of 1628 bp in length
104594 104693: gap of unknown length
104694 106231: contig of 1538 bp in length
106232 106331: gap of unknown length
106332 107693: contig of 1362 bp in length
107694 107793: gap of unknown length
107794 109155: contig of 1362 bp in length
109156 109255: gap of unknown length
109256 110481: contig of 1226 bp in length
110482 110581: gap of unknown length
110582 111881: contig of 1300 bp in length
111882 111981: gap of unknown length
111982 113524: contig of 1543 bp in length
113525 113625: gap of unknown length
113626 114983: contig of 1359 bp in length
114984 115083: gap of unknown length
115084 116145: contig of 1062 bp in length
116146 116245: gap of unknown length
116246 117625: contig of 1380 bp in length
117626 120009: contig of 2284 bp in length
120009 12900: contig of 12900 bp in length

FEATURES
source

1. 246865
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

120010 120109: gap of unknown length
120110 122469: contig of 2360 bp in length
122470 122569: gap of unknown length
122570 124001: contig of 1432 bp in length
124002 124101: gap of unknown length
124102 126352: contig of 2251 bp in length
126353 126452: gap of unknown length
126453 128477: contig of 2025 bp in length
128478 128577: gap of unknown length
128578 130466: contig of 1889 bp in length
130467 130566: gap of unknown length
130567 132505: contig of 1939 bp in length
132506 132605: gap of unknown length
132606 134148: contig of 1543 bp in length
134149 134248: gap of unknown length
134249 136448: contig of 2200 bp in length
136449 136548: gap of unknown length
136549 138255: contig of 1707 bp in length
138256 138355: gap of unknown length
138356 140535: contig of 2180 bp in length
140536 140635: gap of unknown length
140636 142193: contig of 1564 bp in length
142194 142293: gap of unknown length
142294 144906: contig of 2607 bp in length
144907 145006: gap of unknown length
145007 147213: contig of 2207 bp in length
147214 147313: gap of unknown length
147314 148878: contig of 1565 bp in length
148879 148978: gap of unknown length
148979 150906: contig of 1928 bp in length
150907 151006: gap of unknown length
151007 153354: contig of 2348 bp in length
153355 153454: gap of unknown length
153455 155071: contig of 1617 bp in length
155072 155171: gap of unknown length
155172 159010: contig of 3839 bp in length
159011 159110: gap of unknown length
159111 162070: contig of 2960 bp in length
162071 162170: gap of unknown length
162171 164452: contig of 2282 bp in length
164453 164552: gap of unknown length
164553 168294: contig of 3742 bp in length
168295 168394: gap of unknown length
168395 171385: contig of 2991 bp in length
171386 171486: gap of unknown length
171487 175016: contig of 3531 bp in length
175017 175116: gap of unknown length
175117 178766: contig of 3650 bp in length
178767 178866: gap of unknown length
178867 182938: contig of 3972 bp in length
182939 186148: contig of 3210 bp in length
186149 186248: gap of unknown length
186249 190168: contig of 3920 bp in length
190169 190268: gap of unknown length
190269 193524: contig of 3256 bp in length
193525 193624: gap of unknown length
193625 202160: contig of 8536 bp in length
202161 202260: gap of unknown length
202261 210133: contig of 7873 bp in length
210134 210233: gap of unknown length
210234 218141: contig of 7908 bp in length
218142 218241: gap of unknown length
218242 228213: contig of 9972 bp in length
228214 228313: gap of unknown length
228314 237496: contig of 9183 bp in length
237497 237596: gap of unknown length
237597 246865: contig of 9269 bp in length.

Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL

Submitted (16-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIPR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L28659
Center clone name: 36_C_3

* NOTE: This record contains 86 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
649: contig of 649 bp in length
650
749: gap of 100 bp
750
1448: contig of 699 bp in length
1449
1548: gap of 100 bp
1549
2264: contig of 716 bp in length
2265
2364: gap of 100 bp
2365
3078: contig of 714 bp in length
3079
3178: gap of 100 bp
3179
3878: contig of 700 bp in length
3879
3978: gap of 100 bp
3979
4673: contig of 695 bp in length
4674
4773: gap of 100 bp
4774
5453: contig of 680 bp in length
5454
5553: gap of 100 bp
5554
6262: contig of 709 bp in length
6263
7068: contig of 706 bp in length
7069
7168: gap of 100 bp
7169
7864: contig of 696 bp in length
7865
7964: gap of 100 bp
7966
8669: contig of 705 bp in length
8670
8769: gap of 100 bp
8770
9472: contig of 703 bp in length
9473
9573
10268: contig of 696 bp in length
10269
10368: gap of 100 bp
10369
11070: contig of 702 bp in length
11071
11170: gap of 100 bp
11171
11881: contig of 711 bp in length
11882
12689: contig of 708 bp in length
12690
12789: gap of 100 bp
12790
13499: contig of 710 bp in length
13500
13599: gap of 100 bp
13600
14312: contig of 713 bp in length
14313
14413: gap of 100 bp
14414
15124: contig of 712 bp in length
15125
15224: gap of 100 bp
15225
15928: contig of 704 bp in length
15929
16028: gap of 100 bp
16029
16699: contig of 671 bp in length
16700
16799: gap of 100 bp
16800
17493: contig of 694 bp in length
17494
17593: gap of 100 bp

17594
18293
18299: gap of 100 bp
18392: contig of 699 bp in length
18393
19090: contig of 698 bp in length
19091
19190: gap of 100 bp
19191
19391: contig of 711 bp in length
19392
20001: gap of 100 bp
20002
20710: contig of 709 bp in length
20711
20810: gap of 100 bp
20811
21507: contig of 697 bp in length
21508
21607: gap of 100 bp
21608
22314: contig of 707 bp in length
22315
22414: gap of 100 bp
22415
23129: contig of 715 bp in length
23130
23229: gap of 100 bp
23230
23937: contig of 708 bp in length
23938
24037: gap of 100 bp
24038
24744: contig of 707 bp in length
24745
24844: gap of 100 bp
24845
25546: contig of 702 bp in length
25547
25647: gap of 100 bp
25648
26342: contig of 696 bp in length
26343
26442: gap of 100 bp
26443
27153: contig of 711 bp in length
27154
27253: gap of 100 bp
27254
27955: contig of 702 bp in length
27956
28055: gap of 100 bp
28056
28759: contig of 704 bp in length
28760
28859: gap of 100 bp
28860
29569: contig of 710 bp in length
29570
29669: gap of 100 bp
29670
30372: contig of 703 bp in length
30373
30472: gap of 100 bp
30473
31159: contig of 687 bp in length
31160
31259: gap of 100 bp
31260
31973: contig of 714 bp in length
31974
32073: gap of 100 bp
32074
32786: contig of 713 bp in length
32787
32886: gap of 100 bp
32887
33593: contig of 707 bp in length
33594
33693: gap of 100 bp
33694
34394: contig of 701 bp in length
34395
34494: gap of 100 bp
34495
35195: contig of 701 bp in length
35196
35295: gap of 100 bp
35296
35989: contig of 694 bp in length
35990
36089: gap of 100 bp
36090
36785: contig of 696 bp in length
36786
36885: gap of 100 bp
36886
37595: contig of 710 bp in length
37596
37695: gap of 100 bp
37696
38404: contig of 709 bp in length
38405
38504: gap of 100 bp
38505
39215: contig of 711 bp in length
39216
39315: gap of 100 bp
39316
40030: contig of 715 bp in length
40031
40130: gap of 100 bp
40131
40843: contig of 713 bp in length
40844
40943: gap of 100 bp
40944
41659: contig of 716 bp in length
41660
41759: gap of 100 bp
41760
42455: contig of 696 bp in length
42456
42555: gap of 100 bp
42556
43259: contig of 704 bp in length
43260
43359: gap of 100 bp
43360
44060: contig of 701 bp in length
44061
44160: gap of 100 bp
44161
44874: contig of 714 bp in length
44875
44974: gap of 100 bp
44975
45693: contig of 719 bp in length
45694
45793: gap of 100 bp
45794
46498: contig of 705 bp in length
46499
46598: gap of 100 bp
46599
47292: contig of 694 bp in length

Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L28022
Center clone name: 423 F 3

* NOTE: This record contains 86 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

714	813: contig of 713 bp in length
814	813: gap of 100 bp
1543	1543: contig of 730 bp in length
1544	1643: gap of 100 bp
1644	2398: contig of 755 bp in length
2399	2498: gap of 100 bp
2499	3220: contig of 722 bp in length
3221	3220: gap of 100 bp
3321	4071: contig of 751 bp in length
4072	4171: gap of 100 bp
4172	4928: contig of 757 bp in length
4929	5028: gap of 100 bp
5029	5783: contig of 755 bp in length
5784	5883: gap of 100 bp
5884	6640: contig of 757 bp in length
6641	6740: gap of 100 bp
6741	7484: contig of 744 bp in length
7485	7584: gap of 100 bp
7585	8329: contig of 745 bp in length
8330	8429: gap of 100 bp
8430	9147: contig of 718 bp in length
9148	9247: gap of 100 bp
9248	9992: contig of 745 bp in length
9993	10092: gap of 100 bp
10093	10821: contig of 729 bp in length
10822	10921: gap of 100 bp
10923	11666: contig of 745 bp in length
11667	11766: gap of 100 bp
11767	12497: contig of 731 bp in length
12498	12597: gap of 100 bp
12598	13321: contig of 724 bp in length
13322	13421: gap of 100 bp
13422	14166: contig of 745 bp in length
14167	14266: gap of 100 bp
14267	14999: contig of 733 bp in length
15000	15099: gap of 100 bp
15100	15847: contig of 748 bp in length
15848	15947: gap of 100 bp
15948	16693: contig of 746 bp in length
16694	16793: gap of 100 bp
16794	17536: contig of 743 bp in length
17537	17636: gap of 100 bp
17637	18383: contig of 747 bp in length
18384	18483: gap of 100 bp
18484	19222: contig of 739 bp in length
19223	19322: gap of 100 bp
19323	20050: contig of 728 bp in length
20051	20150: gap of 100 bp
20151	20895: contig of 745 bp in length
20896	20995: gap of 100 bp
20996	21747: contig of 752 bp in length
21748	21847: gap of 100 bp
21848	22595: contig of 748 bp in length
22596	22695: gap of 100 bp
22696	23446: contig of 751 bp in length
23447	23546: gap of 100 bp
23547	24299: contig of 753 bp in length

Query Match 23.8%; Score 603.8; DB 2; Length 72409;
Best Local Similarity 83.0%; Pred. No. 6.2e-84;

[illegible]

QY	1884	ATCTCTAAATTCGAATATATTGAGATCAGATAATGGTTGTGGAGACACTAATCC-TTTG	1942
Db	48843	ATCTCTAAATTCGAATATATTGAGATCAGATAATGGTTGTGGAGACACTAATCCTTTTG	48784
QY	1943	ACAATTTGTTGAGGCTTCCTTTTGGAAACCTAAATATGTGCTCAATTTTATATAGACGTTCTGTG	2002
Db	48783	ACAATTTGTTGAGGCTTCCTTTTGGAAACCTAAATATGTGCTCAATTTTATATAGACGTTCTGTG	48724
QY	2003	TTTCTTTTGGGAAAAACATGTAATTCAGTGGTCTTTGGTTTAAATATTTTGTATTTGTATCAT	2062
Db	48723	TTTCTTTTGGGAAAAACATGTAATTCAGTGGTCTTTGGTTTAAATATTTTGTATTTGTATCAT	48664
QY	2063	TAGTTTGAGTTTGCTTAATTTATTTGGCTGAAATCTCCATATTCCTTAATGTCCTCTCAT	2122
Db	48663	TAGTTTGAGTTTGCTTAATTTATTTGGCTGAAATCTCCATATTCCTTAATGTCCTCTCAT	48604
QY	2123	TTTGTCTGCTTCCTTTATTTAAATTTAGAGATAAAATGTTAAATTTATCTCACTCACATATAGTG	2182
Db	48603	TTTGTCTGCTTCCTTTATTTAAATTTAGAGATAAAATGTTAAATTTATCTCACTCACATATAGTG	48544
QY	2183	ATGCTGTTTTTATACATATATATAAATTTATTAATTCATATAATTTATGTTATGTATATA	2242
Db	48543	ATGCTGCTTCTATCTATATATATAAATTTATTAATTCATATAATTTATGTTATGTATATA	48484
QY	2243	TTTGGAGACCTATATCATATATAAATC-----AGAAATTTGTGAAATGACACACATTA	2296
Db	48483	TTTGGAGACCTATTAACATATATACAAACAAATTTAGAAATGTTGATGAAATGACACATTA	48424
QY	2297	TACTTATGATAGTAGCCCTTTTATCTCGTCATAAATGTTATTTGACCTTTTGTCTTAAATTT	2356
Db	48423	TACTTATGATAGTAGCCCTTTTATCTCGTCATAAATGTTATTTGACCTTTTGTCTTAAATTT	48364
QY	2357	TTTTTT-AAATTAATTTGTTTGGTAATCTTTTTCAGCGGTTTATGTCACCTGCTTGCA	2415
Db	48363	TTTTTAAATTAATTTGTTTGGTAATCTTTTTCAGTAGGTTATGTCACCTGCTTGCA	48304
QY	2416	ATTGGTACACAGCTGATTTTATTTAGACATGCTACGCTTTTAAATTTATCTTTTCCAT	2475
Db	48303	ATTGGTACACAGCTGATTTTATTTAGACATGCTACGCTTTTAAATTTATCTTTTCCAT	48244
QY	2476	TTTCAATTTTATTAATTTCTCATATACAAATTTAGTCTCACTTTTACCTTCCCTAGTGTG	2535
Db	48243	TTTCAATTTTATTAATTTCTCATATACAAATTTAGTCTCACTTTTACCTTCCCTAGTGTG	48184
QY	2536	AAATTT 2541	
Db	48183	AAATTT 48178	
RESULT 14			
AC136711		40131 bp	linear
LOCUS			
DEFINITION			
AC136711			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 13.2%; Score 335; DB 6; Length 335;
Best Local Similarity 100.0%; Pred. No. 2.7e-42;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 721 AATTATTTTTCCTTTAACTTTCAACTTCAGGAAACAGTTGGCCCTTGACTCTGTTT 780
DB 1 AATTATTTTTCCTTTAACTTTCAACTTCAGGAAACAGTTGGCCCTTGACTCTGTTT 60

QY 781 GTGGAATAATTTAAACTACTGTTTAAATTTCTTTATTTGTTCTGTAATGACTATTTTACG 840
DB 61 GTGGAATAATTTAAACTACTGTTTAAATTTCTTTATTTGTTCTGTAATGACTATTTTACG 120

QY 841 TCATATAACAATTTTATTTGTTTAAATGACTTTTATTTGTTCTGTAATGACTATTTTAA 900
DB 121 TCATATAACAATTTTATTTGTTTAAATGACTTTTATTTGTTCTGTAATGACTATTTTAA 180

QY 901 TGTCTAGAACAAATTTTATTTGTTCTGATATGACTTTTATTTGTTCTATATGCTATACA 960
DB 181 TGTCTAGAACAAATTTTATTTGTTCTGATATGACTTTTATTTGTTCTATATGCTATACA 240

QY 961 AGATTTTGTGTTGTTTGTGACCGAGTCTTACTCTGTCACCCAGGCTGGAGTGAATGGC 1020
DB 241 AGATTTTGTGTTGTTTGTGACCGAGTCTTACTCTGTCACCCAGGCTGGAGTGAATGGC 300

QY 1021 ATGGTCTCAGCTCACTGCAACCTCCCGCTCCCGGG 1055
DB 301 ATGGTCTCAGCTCACTGCAACCTCCCGCTCCCGGG 335

Search completed: May 21, 2004, 20:49:33
Job time : 10206.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 17:08:16 ; Search time 935.344 Seconds
(without alignments)

10695.346 Million cell updates/sec

Title: US-09-914-651A-1

Perfect score: 335

Sequence: 1 aattatttttctttaacc.....tgcaacctccgctcccg99 335

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rtd:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	93.2	27.8	414	13 BX504814	BX504814 DKFZp686J
C 2	92.8	27.7	684	14 CD642149	CD642149 AGENCOURT
C 3	89.2	26.6	839	14 CD520912	CD520912 AGENCOURT
C 4	89	26.6	1034	13 BX431378	BX431378 BX431378

C 5	88.8	26.5	486	14 CD238898	CD238898 FNPBG09
C 6	88	26.3	867	13 BU561446	BU561446 AGENCOURT
C 7	87.2	26.0	1479	11 BC032458	BC032458 Homo sapi
C 8	86.8	25.9	509	10 BE144987	BE144987 RC2-HT018
C 9	86.6	25.9	609	9 AV760466	AV760466 AV760466
C 10	86	25.7	600	9 AV759518	AV759518 AV759518
C 11	85.8	25.6	398	9 AL696113	AL696113 DKFZp686A
C 12	85.8	25.6	727	14 CD000795	CD000795 AGENCOURT
C 13	85.6	25.5	598	28 AQ053689	AQ053689 RPC111-51
C 14	84.6	25.3	660	9 AL704150	AL704150 DKFZp686D
C 15	84.4	25.2	585	13 BQ778512	BQ778512 i132e06.x
C 16	84	25.1	223	9 AL697921	AL697921 DKFZp686C
C 17	84	25.1	405	14 CD237959	CD237959 FNPAS03
C 18	84	25.1	562	9 AL704374	AL704374 DKFZp686H
C 19	83.8	25.0	683	9 AV733431	AV733431 AV733431
C 20	83.8	25.0	745	13 BU564170	BU564170 AGENCOURT
C 21	83.6	25.0	679	29 AG084405	AG084405 Pan trogl
C 22	83.2	24.8	933	13 BQ897609	BQ897609 AGENCOURT
C 23	83	24.8	822	12 BG575169	BG575169 602598296
C 24	82.8	24.7	213	9 AL712392	AL712392 DKFZp686M
C 25	82.8	24.7	400	9 AV760391	AV760391 AV760391
C 26	82.8	24.7	469	9 AL042753	AL042753 DKFZp434C
C 27	82.8	24.7	689	29 AG123600	AG123600 Pan trogl
C 28	82.8	24.7	1104	11 BC031345	BC031345 Homo sapi
C 29	82.8	24.7	1723	11 BC039095	BC039095 Homo sapi
C 30	82.4	24.6	378	13 BU954284	BU954284 AGENCOURT
C 31	82.4	24.6	722	14 CD639592	CD639592 AGENCOURT
C 32	82.2	24.5	328	13 BX645511	BX645511 DKFZp781F
C 33	82	24.5	485	28 AQ134713	AQ134713 HS 3052.B
C 34	82	24.5	928	13 BX452229	BX452229 BX452229
C 35	81.8	24.4	329	13 BX645928	BX645928 DKFZp781A
C 36	81.8	24.4	362	9 AL712440	AL712440 DKFZp686B
C 37	81.6	24.4	667	29 AG150532	AG150532 Pan trogl
C 38	81.6	24.4	864	10 BF184192	BF184192 601843238
C 39	81.4	24.3	485	28 AQ488128	AQ488128 RPC1-11-2
C 40	81.4	24.3	939	14 CD172331	CD172331 AGENCOURT
C 41	81.2	24.2	210	13 BX476439	BX476439 DKFZp686C
C 42	81.2	24.2	376	14 CD107262	CD107262 AGENCOURT
C 43	81	24.2	537	9 AL157651	AL157651 DKFZp547F
C 44	81	24.2	728	29 AG013357	AG013357 Homo sapi
C 45	80.6	24.1	212	10 AW772536	AW772536 xn04c09.y

ALIGNMENTS

RESULT 1
BX504814/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX504814 414 bp mRNA linear EST 04-SEP-2003
DKFZp686J05182.r1 686 (synonym: hicc3) Homo sapiens cDNA clone
DKFZp686J05182.5', mRNA sequence.

BX504814 GI:32032215

EST.

EST.

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 414)

Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,

Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and

Wiemann,S.

EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,

Wellenreuther,R., et al.)

Unpublished (2003)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by DKFZ (German Cancer Research Center,

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.
No s1 sequence available.
This clone (DKFZp686J05182) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. .414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686J05182"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplEx2; Site_1: SfiII; Site_2: SfiII; cDNA-collection"

FEATURES

source

ORIGIN

Query Match 27.8%; Score 93.2; DB 13; Length 414;
Best Local Similarity 65.7%; Pred. No. 1.2e-07;
Matches 134; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 132 TTTTATTGTTGTTAAAGACATTTATGTTGTCATATGATAAATTTATGTCATAGAAC 191
Db 402 TTTTNTT 343
QY 192 AATTTTATGTCATATGACATTTATGTTATATGTCATATACAACTAGATTTTTTTTG 251
Db 342 TTTTNTT 283
QY 252 TTGTTTTTGAACGAGTCTTACTGTGACCCAGGTGGAGTGTATGGATGCTCAGC 311
Db 282 TTTTCTGAGACAGAGTCTTCTGTGACCCAGGTGGAGTGTATGGCGCATCTCAGC 223
QY 312 TCATGCAACCTCCGCTCCCGG 335
Db 222 TCATGCAACCTCCGCTCCCGG 199

RESULT 2

CD642149/c
LOCUS
DEFINITION
AGENCOURT_14537030 NIH_MGC_191 Homo sapiens cDNA clone
IMAGE:30417982 5', mRNA sequence.
CD642149
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 684)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM220 row: d column: 23
High quality sequence stop: 214.
Location/Qualifiers
1. .684
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES

source

1. .684
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:30417982"
/tissue_type="Pooled"
/lab_host="DH10B (T1 phase-resistant)"
/clone_lib="NIH_MGC_191"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggcgctcgcc); Library is oligo-dT primed
and directionally cloned. FPMC - Peripheral Blood
Mononuclear Cells. RNA was pooled from 3/6hour stimulation
with PMA adn Ionomycin. 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGCCCATTTATGGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGGCACATG-TT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.69
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 27.7%; Score 92.8; DB 14; Length 684;
Best Local Similarity 61.7%; Pred. No. 1.2e-07;
Matches 148; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 93 TTATTGGTTGTAATATGACTATTTTACGTCATATAACAATTTTATTGTTGTTAAATCA 152
Db 242 TTTTNTT 183
QY 153 CTTTATTGTTGTCATATGATAAATTTTATGTCATAGACAAATTTTATTGCTTGATATAT 212
Db 182 TTTTNTT 123
QY 213 GACTTTATTGTTATATGCTATACAACTAGATTTTTCGTTGTTTGTACCGAGTCTTAC 272
Db 122 TTTTNTT 63
QY 273 TCTGTACCCAGCTGGAGTGAATGGCATGCTCAGCTCACTGCAACTCGCCCTCCC 332
Db 62 TCTGTACCCAGCTGGAGTGAATGGCATGCTCAGCTCACTGCAACTCGCCCTCCC 3

RESULT 3

CD520912/c
LOCUS
DEFINITION
AGENCOURT_14356130 NIH_MGC_191 Homo sapiens cDNA clone
IMAGE:30413219 5', mRNA sequence.
CD520912
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 839)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM207 row: n column: 12
High quality sequence stop: 305.
Location/Qualifiers
1. .839

FEATURES

source

1. .839

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30413219"
/tissue_type="Pooled"
/lab_host="PH108 (T1 phage-resistant)"
/clone_lib="NIH_MGC_191"
/notes="Vector: pDNR-LIB; Site 1: SfiI (ggccattagcc); Site 2: SfiI (ggcgccctggcc); Library is oligo-dT primed and directionally cloned. PBMC - Peripheral Blood Mononuclear Cells. RNA was pooled from 3/6hour stimulation with PMA adn Ionomycin. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCGCCATTAGGCC-3' and 3' adaptor sequence: 5'-ATTCAGAGCGCGGCGGCGGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.69 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

```

ORIGIN

Query Match 26.6%; Score 89.2; DB 14; Length 839;
 Best Local Similarity 60.1%; Pred. No. 5.1e-07;
 Matches 163; Conservative 0; Mismatches 106; Indels 2; Gaps 1;

```

QY 65 AAAATTTTAACTACTGTTTAAATTTCTTTATGCTGTAATATGACATTTTACGTCAT 124
Db 296 AANNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 237

QY 125 ATACAAATTTTATGTTTGTAAATGACTTTATTTGTCATATGATATTTATGTC 184
Db 236 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 177

QY 185 ATAGACAATTTTATGTTTGTATATGACTTTATTTGTTATATGCTATACAACTAGAT 244
Db 176 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 117

QY 245 TTTTTCGTTGTTTGACGAGTCTTACTCTGTCCAGGCTGGAGTGAATGGCATGG 304
Db 116 TTTTTCGTTGTTTGACGAGTCTTACTCTGTCCAGGCTGGAGTGAATGGCATGG 58

QY 305 TCTCAGCTCACTGCAACTCCGCTCCCGG 335
Db 57 -ATCAGCTCACTGCAACTCCGCTCCCGG 28

```

RESULT 4

```

BX431378/c
LOCUS
DEFINITION BX431378 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE008YH01
ACCESSION BX431378
VERSION BX431378.1 GI:30776997
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0BAG012ZC11_CS01139_1.
Location/Qualifiers
1. .1034
source

```

FEATURES

Query Match 26.6%; Score 88.8; DB 14; Length 486;
 Best Local Similarity 60.2%; Pred. No. 7.3e-07;
 Matches 147; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

ORIGIN

```

QY 92 TTTATGTTGTTGTAATGACTATTTTACGTATATACAAATTTTATGTTGTTGAATG 151
|||||

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE008YH01"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

```

ORIGIN

Query Match 26.6%; Score 89; DB 13; Length 1034;
 Best Local Similarity 65.2%; Pred. No. 5.2e-07;
 Matches 131; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```

QY 132 TTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTATGTCATAGAAC 191
Db 291 TTTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 232

QY 192 AATTTTATGCTTGATATATGACTTTATTTGTTATATGCTATACAACTAGATTTTGT 251
Db 231 TTTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 172

QY 252 TTGTTTGTGACGAGTCTTACTCTGTCCAGGCTGGAGTGAATGGCATGGTCTCAGC 311
Db 171 GTTTTGTGAGACAGAGTCTTACTCTGTCCAGGCTGGAGTGAATGGCATGGTCTCAGC 112

QY 312 TCAGTCAACTCCGCTCC 332
Db 111 TCAGTCAACTCCGCTCC 91

```

RESULT 5

```

CD238898/c
LOCUS
DEFINITION FNPBGE09 FNP Homo sapiens CDNA, mRNA sequence.
ACCESSION CD238898
VERSION CD238898.1 GI:30982363
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Ma, Y., Qi, X., Zhang, X., Xiao, H., Zhu, Z., Yan, Q., Chen, Z. and Han, Z.
TITLE Fetal Pituitary ESTs
JOURNAL Unpublished (2003)
COMMENT Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn.
Location/Qualifiers
1. .486
source

```

FEATURES

Query Match 26.5%; Score 88.8; DB 14; Length 486;
 Best Local Similarity 60.2%; Pred. No. 7.3e-07;
 Matches 147; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

ORIGIN

```

QY 92 TTTATGTTGTTGTAATGACTATTTTACGTATATACAAATTTTATGTTGTTGAATG 151
|||||

```

[illegible]

RESULT 6	BUS561446	867 bp	linear	EST 16-SEP-2002
LOCUS	BUS561446/c			
DEFINITION	AGNCOURT 10278336 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6592490			
	5', mRNA sequence.			
ACCESSION	BUS561446			
VERSION	BUS61446.1	GI:22911742		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 867)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: CLONTECH cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) cDNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCW2815 row: m column: 02 High quality sequence stop: 343.			

```

FEATURES
  source
    Location/Qualifiers
      1..867
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:6592490"
        /lab_host="DH10B (T1 phage-resistant)"
        /clone_lib="NH_MGC_92"
        /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site:1:
        sfil (ggcgccctcgccg); Site 2: sfil (ggccattatggcc); 5' and
        3' adaptors were used in cloning as follows: 5' adaptor
        sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence:
        5'-ATCTAGAGCGCGGCGGCACATG-dr(30)BN-3' (where B = A,
        C, or G and N = A, C, G, or T). Average insert size
        1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
        inserts by PCR. This library was enriched for full-length
        clones and was constructed by Clontech Laboratories (Palo
        Alto, CA)."
      26.3%; Score 88; DB 13; Length 867;
      Query Match
      Best Local Similarity 59.7%; Pred. No. 8.3e-07;
      Matches 148; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
      Qy 88 TTCTCTTATGGTGTGAATATGACATATTTTCGTCATATAACAATTTTATTGTGTGTA 147
      Db 328 TTTTITTTTTTTTTTTTTTTTTTTTTCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 269

```

Qy	148	AATGAC	TTTATTGTGGTGCATATGATAAAATTTATGTCATAGAACAAATTTTTATGCTGTGA	207
Dβ	268	TT		209
Qy	208	TATATGACTTTA	TGTGTATATGAGCTATACAACCTAGATTTTTTTTGCTTTTTTTCACCGAGT	267
Dβ	208	TT		149
Qy	268	CTTTACTGTGTCACCAGCCTGGAGTGTAAATGGCATGGTCTCAGCTCACTGCAACCTCCGC		327
Dβ	148	GTCCTGGCTGTACCCAGGCCTGGAGTGCAGTGACACCATCTCAGCTTACTACAACTCCGC		89
Qy	328	CTCCCCGGG	335	
Dβ	88	CTCCC	GGG	81

RESULT 7	
BC032458/c	
LOCUS	1479 bp mRNA linear HTC 04-MAR-2003
DEFINITION	Homo sapiens, clone IMAGE:5200744, mRNA.
ACCESSION	BC032458
VERSION	BC032458.1 GI:22749744
KEYWORDS	HTC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
Eukaryota;	
Chordata;	
Cranialia;	
Vertebrata;	
Euteleostomi;	
Mammalia;	
Eutheria;	
Primates;	
Catarrhini;	
Hominidae;	
Homo.	
REFERENCE	1 (bases 1 to 1479) Strausberg R. Direct Submission Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
AUTHORS	
TITLE	
JOURNAL	
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT

Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
cDNA Library Prepared by: National Institutes of Health Intramural
DNA Sequencing by: (NISC),
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter-N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J. C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK file: 64 Row: h Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein

```

FEATURES
  source
    1. 1479
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:520074"
      /tissue_type="brain, adult, 6 pooled whole brains"
      /clone_lib="NIH_MGC_114"
      /lab_host="DH10B"
      /note="Vector: pCMV-SPORT6"

```

```

ORIGIN
Query Match      26.0%; Score 87.2; DB 11; Length 1479;
Best Local Similarity 64.2%; Pred. No. 9.7e-07;
Matches 131; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 132 TTTTATTTGTTGTTAAATGACCTTATTTGTTGTCATATGATAAATTTATGTCATAGAAC 191
Db 1473 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1414
QY 192 AATTTTATGTCATATATGACCTTATTTGTTATATGTCATATGTCATATGTCATATGTCAT 251
Db 1413 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1354
QY 252 TGTGTTTACCGAGCTTCTCTGTCACCGAGCTGAGTGTAATGCGATGCTCAGC 311
Db 1353 TTTTGTGAGTGGAGTCTCACTCTGTCACCGAGCTGAGTGTAATGCGATGCTCAGC 1294
QY 312 TCAGTCAACCTCCGCTCCCGGG 335
Db 1293 TCAGTCAACCTCCAGCTCCCGGG 1270

RESULT 8
LOCUS BE144987 509 bp mRNA linear EST 21-JUN-2000
DEFINITION RC2-HT0187-271099-011-c05 HT0187 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE144987
VERSION BE144987.1 GI:8607711
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 509)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Macukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=RC2-HT0187-271
099-011-c05&t3=1999-10-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 503.
Location/Qualifiers
1..509
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0187"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of

ORIGIN
Query Match      25.9%; Score 86.6; DB 9; Length 609;
Best Local Similarity 60.3%; Pred. No. 1.7e-06;
Matches 161; Conservative 0; Mismatches 104; Indels 2; Gaps 1;

QY 69 TTTTAACTACTGGTTTAAATTTCTTTATTTGTTGTAATGACTATTTTACGTATATAA 128

```

tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

ORIGIN
Query Match      25.9%; Score 86.8; DB 10; Length 509;
Best Local Similarity 59.2%; Pred. No. 1.7e-06;
Matches 148; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 86 AATTTCTTTATTTGGTTGTAATATGACTATTTTACGTATATAAACAATTTTATTTGTTGT 145
Db 181 AAATAAATATCAGTTTATTAAATCTTTTAAAAACACAGCTTTAGGTTTGTTCATTTTC 240
QY 146 TAAATGACTTTATTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTTGCTT 205
Db 241 TCTTCTACTTATCCATTTTCCACTTACTGTTTTCGCTCTTATACATATTAATATGCTTC 300
QY 206 GATATATGACTTTATTTGTTATATGCTATACAACTAGATTTTATTTGTTGTTTTCACCGA 265
Db 301 TGTACTTACTTTGGGTTTATTTGTCATTTTGTGTTTGGTTTGTGTTTGTGTTTGTGAGACA 360
QY 266 GTCTTACTCTGTCACCGAGCTGGAGTGTATGCGATGCTCAGCTCAGCTCAGTCAACCTCC 325
Db 361 GTCTCATTTTGTCAACCGAGCTGGAGTGGTGTGATCTCAACTCACTGACGCTCC 420
QY 326 GCCTCCCGGG 335
Db 421 GCCTCCCGGG 430

```

```

RESULT 9
LOCUS AV760466 609 bp mRNA linear EST 19-OCT-2000
DEFINITION AV760466 MDS Homo sapiens cDNA clone MDSBKFL1 5', mRNA sequence.
ACCESSION AV760466
VERSION AV760466.1 GI:10918314
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 609)
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,
Lu, G., Yang, Y., Gao, G., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA MDS clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1..609
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MDSBKFL1"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="MDS"
/note="Vector: pTriplex2; Site 1: sfiI; Site 2: sfiIB"

```

FEATURES
source

ORIGIN

Query Match 25.9%; Score 86.6; DB 9; Length 609;
Best Local Similarity 60.3%; Pred. No. 1.7e-06;
Matches 161; Conservative 0; Mismatches 104; Indels 2; Gaps 1;

QY 69 TTTTAACTACTGGTTTAAATTTCTTTATTTGTTGTAATGACTATTTTACGTATATAA 128

[illegible]

RESULT 12
 CD000795
 LOCUS
 DEFINITION
 AGENCOURT 13645840 NIH_MGC 186 Homo sapiens cDNA clone
 IMAGE:30321777 5', mRNA sequence.
 CD000795
 VERSION
 CD000795.1 GI:30295326
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 727)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Falkovits
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: NDCM121 row: 1 column: 10
 High quality sequence stop: 528.
 Location/Qualifiers
 1..727
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30321777"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC 186"
 /note="Organ: Pooled Skin; Vector: pDNR-LIB; Site 1: SfiI
 (ggccattagggc); Site 2: SfiI (ggccctctggcc); Library is
 oligo-dT primed and directionally cloned. cDNA was
 prepared from a pooled samples of tissues from Skin,
 meninges, duramater, pia matter and choroid plexus. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGGAGCGGCATG-dt(30)BN-3'
 (where B = A, C, G or T). Average
 insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
 Library"

FEATURES

source

CD000795 727 bp mRNA linear EST 01-MAY-2003
 AGENCOURT 13645840 NIH_MGC 186 Homo sapiens cDNA clone
 IMAGE:30321777 5', mRNA sequence.
 CD000795
 VERSION
 CD000795.1 GI:30295326
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 727)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Falkovits
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: NDCM121 row: 1 column: 10
 High quality sequence stop: 528.
 Location/Qualifiers
 1..727
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30321777"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC 186"
 /note="Organ: Pooled Skin; Vector: pDNR-LIB; Site 1: SfiI
 (ggccattagggc); Site 2: SfiI (ggccctctggcc); Library is
 oligo-dT primed and directionally cloned. cDNA was
 prepared from a pooled samples of tissues from Skin,
 meninges, duramater, pia matter and choroid plexus. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGGAGCGGCATG-dt(30)BN-3'
 (where B = A, C, G or T). Average
 insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
 Library"

ORIGIN

Query Match 25.6%; Score 85.8; DB 14; Length 727;
 Best Local Similarity 59.0%; Pred. No. 2.2e-06;
 Matches 147; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
 QY 87 ATTCTTTATGGTGTAAATGACTATTATTACGTCATATAACAATTTTATGTTTGT 146
 DB 159 ATTAAATTATAGAACAGACTAAATTTTATAATTTATTTAATATATATATACATAT 218
 QY 147 AAATGACTTTATGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATGCTTG 206
 DB 219 ATATGTTGTAATATATACATATATATATGTTGTTATATATATATATATATATGTT 278
 QY 207 ATATATGACTTTTATGTTTATATGGCTATACAACTAGATTTTGTGTTTATGCTTG 266
 DB 279 GGTGTAATATATGATGTATATATATATATATATATATATATATATATATATAT 338
 QY 267 TCTTACTCTGTCAACCGAGCTGGAGTGAATGGCATGGTCTCAGCTCACTGCAACTCG 326
 DB 339 TCTCACTCTGTCAACCGAGCTGGAGTGAATGGCATGGTCTCAGCTCACTGCAACTCTG 398

QY 327 CCTCCCGG 335
 DB 399 CCTCCCGG 407
 RESULT 13
 AQ053689
 LOCUS
 DEFINITION
 AQ053689 598 bp DNA linear GSS 20-APR-1999
 RPCI11-51010 TK RPCI-11 Homo sapiens genomic clone RPCI-11-51010,
 genomic survey sequence.
 ACCESSION
 AQ053689
 VERSION
 AQ053689.1 GI:3348612
 KEYWORDS
 GSS.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 598)
 Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,
 Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and
 Venter, J.C.
 Use of BAC End Sequences for Sequence-Ready Map Building (1998)
 Unpublished (1998)
 Other GSSs: RPCI11-51010.TJ
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadam@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
 Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Class: BAC ends.

FEATURES

source

Location/Qualifiers
 1..598
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7519545"
 /db_xref="taxon:9606"
 /clone="RPCI-11-51010"
 /sex="Male"
 /cell_type="Lymphocytes"
 /clone_lib="RPCI-11"
 /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPCI11 Human Male BAC Library"

ORIGIN

Query Match 25.6%; Score 85.6; DB 28; Length 598;
 Best Local Similarity 60.2%; Pred. No. 2.6e-06;
 Matches 142; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
 QY 100 TTGTAATGACTATTATTACGTCATATAACAATTTTATGTTTGTAAATGACTTTAT 159
 DB 286 TTTAAACATTTAATTTACATACAATTTTAAATATTTTAAAAAGTTATATATATATA 345
 QY 160 GTTTCATATGATAATTTTATGTCATAGAACAAATTTTATGTTTGTGATATGACTTTA 219
 DB 346 TACATA 405
 QY 220 TTGTTATAGGCTATACAACTAGATTTTGTGTTTGTGACCGAGTCTTACTCTGTCA 279
 DB 406 TACATATATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 465
 QY 280 CCCAGCTGGAGTGTATGGCAGTCTCAGCTCACTCAACCTCCGCTCCCGG 335
 DB 466 CCCAGCTGGAGTGTATGGCAGTCTCAGCTCACTCAACCTCCGCTCCCGG 521

RESULT 14
AL704150/c
LOCUS
DEFINITION
AL704150
ACCESSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

660 bp mRNA linear EST 04-SEP-2003
DKFZp686D0529_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DKFZp686D0529_5', mRNA sequence.
AL704150.1 GI:19687505
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 660)
Lauber, J., Bahr, A., Wewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
EST (Lauber, J., Bahr, A., Wewes, H.W., Weil, B., Amid, C., et al.)
Unpublished (2003)
Contact: MIPS
MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No si sequence available.
This clone (DKFZp686D0529) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. .660
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686D0529"
/lab_host="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match 25.3%; Score 84.6; DB 9; Length 660;
Best Local Similarity 68.4%; Pred. No. 3.9e-06;
Matches 117; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 165 TCATATGATATTTTATGCTATGACAAATTTTATGCTGATATATGCTTATGTT 224
Db 427 TCATGGCATAATGATTTTATATATATATATATGCTGCTGCTGCTGCTATAT 368
QY 225 ATATGCTATACAACTAGATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 284
Db 367 ATATGCTATACAACTAGATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 308
QY 285 GCTGGAGTGAATGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 335
Db 307 GCTGGAGTGAATGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257

RESULT 15
BQ778512
LOCUS
DEFINITION
BQ778512
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

585 bp mRNA linear EST 26-JUL-2002
il32e06.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6031595 3',
mRNA sequence.
GI:21986984
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 585)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

TITLE
JOURNAL
COMMENT

Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blisstein, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R.,
Williams, T., Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: il32e06.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -400P from Gibco
High quality sequence stop: 447.

FEATURES
source

1. .585
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6031595"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming. ~1kb. 5'
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match 25.2%; Score 84.4; DB 13; Length 585;
Best Local Similarity 66.5%; Pred. No. 4.4e-06;
Matches 121; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 154 TTTATGTTGTCATATGATATTTTATGCTATGACAAATTTTATGCTGATATATG 213
Db 18 TTGACATTTATTAATGCAAGACACATTTGATGACAAATATTCATTCAGGATATGC 77
QY 214 ACTTTATGTTATATGGCTATACACTAGATTTTATGTTGTTGTTGTTGTTGTT 273
Db 78 CTTCATATCCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 137
QY 274 CTGTCACCCAGGCTGAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 333
Db 138 CTTCACCCAGGCTGAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 197
QY 334 GG 335
Db 198 GG 199

Search completed: May 21, 2004, 23:03:19
Job time : 939.344 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 20:49:36 ; Search time 151.193 Seconds
(without alignments)
10069.828 Million cell updates/sec

Title: US-09-914-651A-1

Perfect score: 335

Sequence: 1 aattatttttcttaacc.....tgcacacctccgctcccggg 335

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	312	93.1	335	16	US-10-431-791-3
2	262	78.2	454	16	US-10-431-791-1
3	246	73.4	257	16	US-10-431-791-18
4	94	28.1	13444	9	US-09-764-877-3660
5	94	28.1	13444	16	US-10-242-515-3660
6	89.8	26.8	53332	9	US-09-801-861-3
7	89.8	26.8	53332	15	US-10-224-562-3
8	86.4	25.8	40491	13	US-10-087-192-1426
9	85.4	25.5	567	13	US-10-027-632-138568
10	85.4	25.5	567	13	US-10-027-632-138569
11	85.4	25.5	567	13	US-10-027-632-138570
12	85.4	25.5	567	16	US-10-027-632-138568
13	85.4	25.5	567	16	US-10-027-632-138569
14	85.4	25.5	567	16	US-10-027-632-138570

```

15 85.2 25.4 2185 16 US-10-074-024-821
16 85 25.4 237326 15 US-10-301-844-2
17 85 25.4 378361 10 US-09-901-136-3
18 84.8 25.3 22197 13 US-10-087-192-1156
19 84.6 25.3 133632 13 US-10-087-192-1810
20 83.8 25.0 72604 15 US-10-162-497-7
21 83.6 25.0 18998 9 US-09-764-869-1952
22 83.6 25.0 18998 15 US-10-091-504-1952
23 83.6 25.0 18998 16 US-10-227-577-1952
24 83.6 25.0 108316 15 US-10-232-798-1789
25 83.6 25.0 108317 15 US-10-017-161-2143
26 83.4 24.9 1108 13 US-10-027-632-118828
27 83.4 24.9 1108 16 US-10-027-632-118828
28 83.2 24.8 650 13 US-10-027-632-214024
29 83.2 24.8 650 16 US-10-027-632-214024
30 82.8 24.7 3015 13 US-10-027-632-116383
31 82.8 24.7 3015 13 US-10-027-632-116383
32 82.8 24.7 3015 13 US-10-027-632-116384
33 82.8 24.7 3015 16 US-10-027-632-116385
34 82.8 24.7 3015 16 US-10-027-632-116383
35 82.8 24.7 3015 16 US-10-027-632-116384
36 82.6 24.7 762 13 US-10-027-632-116385
37 82.6 24.7 762 16 US-10-027-632-163521
38 82.4 24.6 1119 13 US-10-027-632-163521
39 82.4 24.6 1119 13 US-10-027-632-117192
40 82.4 24.6 1119 16 US-10-027-632-117192
41 82.4 24.6 1119 16 US-10-027-632-117192
42 82.4 24.6 147309 9 US-09-742-312-3
43 82.4 24.6 147309 15 US-10-436-185-3
44 82.2 24.5 687 13 US-10-027-632-27766
45 82.2 24.5 687 16 US-10-027-632-27766

```

ALIGNMENTS

```

RESULT 1
US-10-431-791-3
; Sequence 3, Application US/10431791
; Publication No. US20030235874A1
; GENERAL INFORMATION:
; APPLICANT: Kao, Chinghai
; APPLICANT: Lee, Sang-Jin
; APPLICANT: Kim, Hong-Sup
; APPLICANT: Lee, KangRyul
; APPLICANT: Yu, Rong
; TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and
; FILE REFERENCE: Methods of Use Thereof
; FILE REFERENCE: 1857-ARTI.0222US
; CURRENT APPLICATION NUMBER: US/10/431,791
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: 60/378,920
; PRIOR FILING DATE: 2002-05-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-10-431-791-3

```

Query Match 93.1%; Score 312; DB 16; Length 335;
Best Local Similarity 99.4%; Pred. No. 5.9e-53;
Matches 334; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```

QY 1 AATTATTTTTCCTTTTAACTTCAACCTTCAAGAAACCACTGCGCTGCTGTTT 60
Db 1 AATTATTTTTCCTTTTAACTTCAACCTTCAAGAAACCACTGCGCTGCTGTTT 60
QY 61 GTGGAATTTTAACTTCACTGCTTTAACTTTCTTTTATTTGTTTAACTATATACCTATTTTACG 120

```

Db 61 GTGGAATAATTTAAACTACTGGTTTAAATTTCTTTATTGGTTTGTAAATGACTATTTTACG 120

Qy 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTAT-TGTTGTCATATGATAAATTTT 179

Db 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTTGTTGTCATATGATAAATTTT 180

Qy 180 ATGTCATAGAACAAATTTTATTGCTTGATATATGACTTTTATTGTTTATATGCTATACAAAC 239

Db 181 ATGTCATAGAACAAATTTTATTGCTTGATATATGACTTTTATTGTTTAT-TGGCTATACAAAC 239

Qy 240 TAGATTTTGTGTTTGTGACCGAGTCTACTCTGTCACCCAGGCTGGAGTGAATGG 299

Db 240 TAGATTTTGTGTTTGTGACCGAGTCTACTCTGTCACCCAGGCTGGAGTGAATGG 299

Qy 300 CATGCTCTCAGCTCACTCAACCTCCGCTCCCGGG 335

Db 300 CATGCTCTCAGCTCACTCAACCTCCGCTCCCGGG 335

RESULT 2

US-10-431-791-1

Sequence 1, Application US/10431791

Publication No. US20030235874A1

GENERAL INFORMATION:

APPLICANT: Kao, Chinghai

APPLICANT: Lee, Sang-Jin

APPLICANT: Kim, Hong-Sup

APPLICANT: Lee, Kangryul

APPLICANT: Yu, Rong

TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and

FILE REFERENCE: 1857-ARTI.0222US

CURRENT APPLICATION NUMBER: US/10/431,791

CURRENT FILING DATE: 2003-05-08

PRIOR APPLICATION NUMBER: 60/378,920

PRIOR FILING DATE: 2002-05-08

NUMBER OF SEQ ID NOS: 18

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 454

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic Sequence

US-10-431-791-1

Query Match 78.2%; Score 262; DB 16; Length 454;

Best Local Similarity 100.0%; Pred. No. 6.1e-43; Mismatches 0; Indels 0; Gaps 0;

Matches 262; Conservative 0;

Qy 1 AATTATTTTTCCTTTAAACCTTTCAAACTCAAGGAAACCCAGTTGGCTTGACTCTGTTT 60

Db 193 AATTATTTTTCCTTTAAACCTTTCAAACTCAAGGAAACCCAGTTGGCTTGACTCTGTTT 252

Qy 61 GTGGAATAATTTAAACTACTGGTTTAAATTTCTTTATTGGTTGTAATGACTATTTTACG 120

Db 253 GTGGAATAATTTAAACTACTGGTTTAAATTTCTTTATTGGTTGTAATGACTATTTTACG 312

Qy 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTTGTTGTCATATGATAAATTTA 180

Db 313 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTTGTTGTCATATGATAAATTTA 372

Qy 181 TGTCAATAGAACAAATTTTATTGCTTGATATATGACTTTTATTGTTTATATGCTATACAACT 240

Db 373 TGTCAATAGAACAAATTTTATTGCTTGATATATGACTTTTATTGTTTATATGCTATACAACT 432

Qy 241 AGATTTTGTGTTTGTGAC 262

Db 433 AGATTTTGTGTTTGTGAC 454

RESULT 3

US-10-431-791-18

Sequence 18, Application US/10431791

Publication No. US20030235874A1

GENERAL INFORMATION:

APPLICANT: Kao, Chinghai

APPLICANT: Lee, Sang-Jin

APPLICANT: Kim, Hong-Sup

APPLICANT: Lee, Kangryul

APPLICANT: Yu, Rong

TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and

FILE REFERENCE: 1857-ARTI.0222US

CURRENT APPLICATION NUMBER: US/10/431,791

CURRENT FILING DATE: 2003-05-08

PRIOR APPLICATION NUMBER: 60/378,920

PRIOR FILING DATE: 2002-05-08

NUMBER OF SEQ ID NOS: 18

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 18

LENGTH: 257

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic Sequence

US-10-431-791-18

Query Match 73.4%; Score 246; DB 16; Length 257;

Best Local Similarity 99.6%; Pred. No. 7.7e-40; Mismatches 0; Indels 1; Gaps 1;

Matches 257; Conservative 0;

Qy 1 AATTATTTTTCCTTTAAACCTTTCAAACTCAAGGAAACCCAGTTGGCTTGACTCTGTTT 60

Db 1 AATTATTTTTCCTTTAAACCTTTCAAACTCAAGGAAACCCAGTTGGCTTGACTCTGTTT 60

Qy 61 GTGGAATAATTTAAACTACTGGTTTAAATTTCTTTATTGGTTGTAATGACTATTTTACG 120

Db 61 GTGGAATAATTTAAACTACTGGTTTAAATTTCTTTATTGGTTGTAATGACTATTTTACG 120

Qy 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTTGTTGTCATATGATAAATTTA 180

Db 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTTGTTGTCATATGATAAATTTA 180

Qy 181 TGTCAATAGAACAAATTTTATTGCTTGATATATGACTTTTATTGTTTATATGCTATACAACT 240

Db 181 TGTCAATAGAACAAATTTTATTGCTTGATATATGACTTTTATTGTTTAT-TGGCTATACAACT 239

Qy 241 AGATTTTGTGTTTGT 258

Db 240 AGATTTTGTGTTTGT 257

RESULT 4

US-09-764-877-3660/c

Sequence 3660, Application US/09764877

Patent No. US20020147140A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC005

CURRENT APPLICATION NUMBER: US/09/764,877

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 4031

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3660

LENGTH: 13444

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-877-3660

Query Match 28.1%; Score 94; DB 9; Length 13444;

Best Local Similarity 65.0%; Pred. No. 6.6e-09; Mismatches 75; Indels 0; Gaps 0;

Matches 139; Conservative 0;


```

; ORGANISM: Human
US-10-027-632-138569

Query Match      25.5%; Score 85.4; DB 13; Length 567;
Best Local Similarity 65.4%; Pred. No. 1.1e-07;
Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 145 TTAATGACTTTATGTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTGCT 204
Db 298 TAAAAAAATTCATTTGGATGGCAATTTATAGCATCTTAGGTGAACAAAGAAATTTGTT 357
QY 205 TGATATATGACTTTATGTTATATGCTATACACTAGATTTTTTGTGTTTGGACCG 264
Db 358 TGTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTGGATGG 417
QY 265 AGTCTTACTCTGCACCCAGGCTGGAGTGAATGGCATGTTCTCAGCTCACTGCAACCTC 324
Db 418 AGCTTGTCTGTCTCCAGGCTGGAGTGCAGTGGCAACATCTCAGCTCACTGCAAACTC 477
QY 325 CGCCTCCGGG 335
Db 478 CACCTTCGGG 488

RESULT 11
US-10-027-632-138570
; Sequence 138570, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138570
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-138570

Query Match      25.5%; Score 85.4; DB 13; Length 567;
Best Local Similarity 65.4%; Pred. No. 1.1e-07;
Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 145 TTAATGACTTTATGTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTGCT 204
Db 298 TAAAAAAATTCATTTGGATGGCAATTTATAGCATCTTAGGTGAACAAAGAAATTTGTT 357
QY 205 TGATATATGACTTTATGTTATATGCTATACACTAGATTTTTTGTGTTTGGACCG 264
Db 358 TGTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTGGATGG 417
QY 265 AGTCTTACTCTGCACCCAGGCTGGAGTGAATGGCATGTTCTCAGCTCACTGCAACCTC 324
Db 418 AGCTTGTCTGTCTCCAGGCTGGAGTGCAGTGGCAACATCTCAGCTCACTGCAAACTC 477
QY 325 CGCCTCCGGG 335
Db 478 CACCTTCGGG 488

RESULT 12
US-10-027-632-138568
; Sequence 138568, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138568
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-138568

Query Match      25.5%; Score 85.4; DB 16; Length 567;
Best Local Similarity 65.4%; Pred. No. 1.1e-07;
Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 145 TTAATGACTTTATGTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTGCT 204
Db 298 TAAAAAAATTCATTTGGATGGCAATTTATAGCATCTTAGGTGAACAAAGAAATTTGTT 357
QY 205 TGATATATGACTTTATGTTATATGCTATACACTAGATTTTTTGTGTTTGGACCG 264
Db 358 TGTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTGGATGG 417
QY 265 AGTCTTACTCTGCACCCAGGCTGGAGTGAATGGCATGTTCTCAGCTCACTGCAACCTC 324
Db 418 AGCTTGTCTGTCTCCAGGCTGGAGTGCAGTGGCAACATCTCAGCTCACTGCAAACTC 477
QY 325 CGCCTCCGGG 335
Db 478 CACCTTCGGG 488

RESULT 13
US-10-027-632-138569
; Sequence 138569, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```



```

; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138569
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-138569

Query Match      25.5%; Score 85.4; DB 16; Length 567;
Best Local Similarity 65.4%; Pred. No. 1.1e-07;
Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 145 TTAATGACATTTATTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTGCT 204
Db 298 TAAAAAAATTCATTGGATGGCAATTTATAGCATCTTAGGTGAACAAAAGAAATTTGCT 357
QY 205 TGATATATGACATTTATTTGTCATATGCTATACAACTAGATTTTGTGTTTGGACCG 264
Db 358 TGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 417
QY 265 AGCTTACTCTGTCACCCAGGCTGGAGTGAATGTCATGCTCAGCTCAGTCAACCTC 324
Db 418 AGCTTGTCTGTCCTCCAGGCTGGAGTGCATGTCGACCAATCTCAGCTCAGTCAACCTC 477
QY 325 CGCCTCCCGGG 335
Db 478 CACCTTCCGGG 488

RESULT 14
US-10-027-632-138570
; Sequence 138570, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138570
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-138570

```

```

Query Match      25.5%; Score 85.4; DB 16; Length 567;
Best Local Similarity 65.4%; Pred. No. 1.1e-07;
Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 145 TTAATGACATTTATTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTGCT 204
Db 298 TAAAAAAATTCATTGGATGGCAATTTATAGCATCTTAGGTGAACAAAAGAAATTTGCT 357
QY 205 TGATATATGACATTTATTTGTCATATGCTATACAACTAGATTTTGTGTTTGGACCG 264
Db 358 TGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 417
QY 265 AGCTTACTCTGTCACCCAGGCTGGAGTGAATGTCATGCTCAGCTCAGTCAACCTC 324
Db 418 AGCTTGTCTGTCCTCCAGGCTGGAGTGCATGTCGACCAATCTCAGCTCAGTCAACCTC 477
QY 325 CGCCTCCCGGG 335
Db 478 CACCTTCCGGG 488

RESULT 15
US-10-074-024-821
; Sequence 821, Application US/10074024
; Publication No. US20030232975A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC001C1
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US/10/074,024
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 879
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 821
; LENGTH: 2185
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-074-024-821

Query Match      25.4%; Score 85.2; DB 16; Length 2185;
Best Local Similarity 66.1%; Pred. No. 1.9e-07;
Matches 123; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 150 TGACTTTATTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTGCTTGATA 209
Db 365 TAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 424
QY 210 TATGACTTTTATTTGTCATATGCTATACAACTAGATTTTGTGTTTGGACCGAGTCT 269
Db 425 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 484
QY 270 TACTCTGTCAACCCAGGCTGGAGTGAATGTCATGCTCAGCTCAGTCAACCTCCGCCT 329
Db 485 TACTCTGTCAACCCAGGCTGGAGTGAATGTCATGCTCAGCTCAGTCAACCTCCGCCT 544
QY 330 CCGGG 335
Db 545 CCGGG 550

Search completed: May 22, 2004, 02:28:08
Job time : 154.193 secs

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 17:11:06 ; Search time 25.8588 Seconds
(without alignments)
7189.364 Million cell updates/sec

Title: US-09-914-651A-1
Perfect score: 335
Sequence: 1 aattatttttccttaacc.....tgcacacctccgctcccggg 335

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89.8	26.8	53332	4	US-09-801-861-3
C 2	83.8	25.0	72604	4	US-09-268-992-7
C 3	83.8	25.0	72604	4	US-09-657-474-7
C 4	83.2	24.8	28720	4	US-09-341-587-7
C 5	81.8	24.4	62804	4	US-09-800-960-3
C 6	81.8	24.4	62804	4	US-10-096-960-3
C 7	81.6	24.4	6799	4	US-09-620-312D-299
C 8	79.8	23.8	28001	4	US-09-819-993-3
C 9	79.8	23.8	28001	4	US-10-193-295-3
C 10	79	23.6	36159	4	US-09-749-588-3
C 11	78.6	23.5	11811	3	US-09-078-294-7
C 12	77.8	23.2	246240	2	US-08-724-394A-20
C 13	77.8	23.2	246240	2	US-08-724-394A-21
C 14	77.8	23.2	246240	2	US-08-724-394A-22
C 15	77.4	23.1	29629	4	US-09-729-995-3
C 16	77.4	23.1	29629	4	US-10-135-689-3
C 17	77.2	23.0	49312	4	US-09-671-317-485
C 18	77.2	23.0	392000	4	US-10-027-983-11
C 19	77	23.0	12597	4	US-09-705-299-12
C 20	76.8	22.9	3701	4	US-09-220-132-57
C 21	76.8	22.9	84495	4	US-09-797-906-3
C 22	76.6	22.9	38844	4	US-09-734-675-3
C 23	76.2	22.7	4428	4	US-09-023-655-1109
C 24	76.2	22.7	6038	3	US-09-305-639-4
C 25	76.2	22.7	6038	4	US-09-525-160B-2
C 26	76.2	22.7	7622	3	US-09-305-639-1
C 27	76.2	22.7	7622	4	US-09-525-160B-1

28 76 22.7 45546 4 US-09-146-053-6 Sequence 6, Appli
29 75.8 22.6 6405 4 US-09-281-481A-18 Sequence 18, Appl
30 75.8 22.6 13158 2 US-08-687-080-105 Sequence 105, Appl
c 31 75.2 22.4 99500 4 US-09-798-096-10 Sequence 10, Appl
32 75 22.4 80246 3 US-09-078-294-4 Sequence 4, Appli
33 74.6 22.3 162450 4 US-09-345-882-1 Sequence 1, Appli
34 74.4 22.2 1856 1 US-08-157-171-3 Sequence 3, Appli
35 74.4 22.2 1856 3 US-09-050-159-128 Sequence 128, Appl
36 74.4 22.2 72928 3 US-09-009-913-1 Sequence 1, Appli
c 37 74.4 22.2 83450 4 US-09-811-469-3 Sequence 3, Appli
38 74.4 22.2 116592 4 US-09-818-512-3 Sequence 3, Appli
c 39 74.2 22.1 63000 4 US-09-780-172-18 Sequence 18, Appl
c 40 73.8 22.0 398 4 US-09-621-976-10198 Sequence 10198, A
41 73.8 22.0 9365 4 US-09-608-285A-8 Sequence 8, Appli
42 73.8 22.0 9365 4 US-09-350-836B-8 Sequence 8, Appli
43 73.8 22.0 9365 4 US-09-370-285-8 Sequence 8, Appli
44 73.8 22.0 9365 4 US-09-557-800C-8 Sequence 8, Appli
45 73.8 22.0 9365 4 US-09-370-625A-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-801-861-3
; Sequence 3, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-3

Query Match 26.8%; Score 89.8; DB 4; Length 53332;
Best Local Similarity 59.8%; Pred. No. 2.5e-12;
Matches 168; Conservative 0; Mismatches 112; Indels 1; Gaps 1;
Qy 54 TCTGTTGTGGAAATTTTAAACTACTGGTTTAAATTTCTTTTATGTTGTAATGACTA 113
Db 31467 TATATATTTATATATATTTTATATATATATATATATATATATATATATATATATA 31526
Qy 114 TTTTACCTCATATACAAATTTTATGTTGTTAAATGACTTTATGTTGTCATATGAT 173
Db 31527 TATATTTATTA 31586
Qy 174 AATTTTATGTCATAGAACAAATTTTATGTTGTCATATATGACTTTATGTTTATGCTA 233
Db 31587 TATATTTATATATATTTTATATATATATATATATATATATATATATATATATATA 31645
Qy 234 TACAACATAGATTTTTTTTGTGTTTGTGACCGAGTCTTCTGTCACCCAGGCTGGAGTG 293
Db 31646 TATATATATCTATTTTTTTTTTTTGTGAGATGAGTCTCACTCTATGCGCAGGCTGGAGTG 31705
Qy 294 TAATGCGATGCTCAGCTCAGTCACTGCAACCTCCCGCTCCCGG 334
Db 31706 CAGTGGCAGATCTCAGCTCAGTCACTGCAACCTCCCGCTCCCGG 31746

RESULT 2
US-09-268-992-7/c
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:


```
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match      23.2%; Score 77.8; DB 2; Length 246240;
Best Local Similarity 59.7%; Pred. No. 2.3e-09;
Matches 151; Conservative 0; Mismatches 97; Indels 5; Gaps 1;

QY 83 TTTAAATCTCTTATGTTGTTAATATGACTATTTTACGTCATATAACAATTTTATGTT 142
Db 179076 TCTAAATTTGTTAATGCTCTCTTTTAAATGCAAAAGACACATCACTTTACCTTCTTCA 179135

QY 143 TGTAAATGACCTTATTTGTTGTCATATGATATAATTTTATGTCATAGAACAAATTTTATG 202
Db 179136 TTACCAAGTGCCTTTGGAATTTTCTGTCATCTAATTTCTGTTTCATTTATTCATATAT 179195

QY 203 CTGTGATATAGACTTTTATGTTTATATGCTATACAACTAGATTTTGTGTTGTTGAC 262
Db 179196 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 179250

QY 263 CGAGTCTTACTCTGTCCACCCAGGCTGGAGTGAATGCGCATGGTCTCAGCTCACTGCAACC 322
Db 179251 AGAGTCTCACACTGTCCACCCAGGCTGGAGTGAATGCGCATGGTCTCAGCTCACTGCAACC 179310

QY 323 TCCGCTCCCGGG 335
Db 179311 TCTGCTCTCTGAG 179323

RESULT 14
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant

; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22

Query Match      23.2%; Score 77.8; DB 2; Length 246240;
Best Local Similarity 59.7%; Pred. No. 2.3e-09;
Matches 151; Conservative 0; Mismatches 97; Indels 5; Gaps 1;

QY 83 TTTAAATCTCTTATGTTGTTAATATGACTATTTTACGTCATATAACAATTTTATGTT 142
Db 179076 TCTAAATTTGTTAATGCTCTCTTTTAAATGCAAAAGACACATCACTTTACCTTCTTCA 179135

QY 143 TGTAAATGACCTTATTTGTTGTCATATGATATAATTTTATGTCATAGAACAAATTTTATG 202
Db 179136 TTACCAAGTGCCTTTGGAATTTTCTGTCATCTAATTTCTGTTTCATTTATTCATATAT 179195

QY 203 CTGTGATATAGACTTTTATGTTTATATGCTATACAACTAGATTTTGTGTTGTTGAC 262
Db 179196 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 179250

QY 263 CGAGTCTTACTCTGTCCACCCAGGCTGGAGTGAATGCGCATGGTCTCAGCTCACTGCAACC 322
Db 179251 AGAGTCTCACACTGTCCACCCAGGCTGGAGTGAATGCGCATGGTCTCAGCTCACTGCAACC 179310

QY 323 TCCGCTCCCGGG 335
Db 179311 TCTGCTCTCTGAG 179323

RESULT 15
US-09-729-995-3/c
; Sequence 3, Application US/09729995
; Patent No. 6436206
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904
; CURRENT APPLICATION NUMBER: US/09/729,995
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 29629
; TYPE: DNA
; ORGANISM: Human
; US-09-729-995-3

Query Match      23.1%; Score 77.4; DB 4; Length 29629;
Best Local Similarity 71.3%; Pred. No. 1.9e-09;
Matches 102; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 193 ATTTTATGCTTGATATATGACTTTTATGTTTATATGCTTATGCTATACAACTAGATTTTGT 252
Db 17380 ATTTTCTTTTGTGTTGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 17321

QY 253 TGTGTTTGGCCGAGTCTTACTCTGTCACCCAGGCTGGAGTGAATGCGCATGGTCTCAGCT 312
Db 17320 TTTTTCAGACAGAGTCTGCTGTCGCCAGGCTGGAGTGAATGCGCATGGTCTCAGCT 17261

QY 313 CACTGCAACCTCCGCTCCCGGG 335
Db 17260 CACTGAAACCTCTGCTCCCGGG 17238

Search completed: May 21, 2004, 23:07:07
Job time : 28.8588 secs
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 16:24:44 ; Search time 138.729 Seconds
(without alignments)
10258.453 Million cell updates/sec

Title: US-09-914-651A-1

Perfect score: 335

Sequence: 1 aattatttttctttaaacc.....tgcacacctcgccctccggg 335

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	335	100.0	335	3 AAA94789	Aaa94789 Human Pro
2	94	28.1	13444	4 AAL37295	Aal37295 Human mus
3	94	28.1	13444	7 ABX60283	Abx60283 cDNA enco
4	90.6	27.0	18252	4 AAK87554	Aak87554 Human imm
5	90.2	26.9	39110	4 AAK73087	Aak73087 Human imm
6	90.2	26.9	39110	4 AAK71825	Aak71825 Human imm
7	90.2	26.9	39110	4 AAK87555	Aak87555 Human imm
8	90.2	26.9	39110	7 ABZ68185	Abz68185 Human sec
9	90.2	26.9	39110	7 ABZ74663	Abz74663 Secreted
10	89.8	26.8	53332	6 AAL48890	Aal48890 Human Pft
11	87.2	26.0	140167	6 ABT10146	Abt10146 Human bre
12	86.4	25.8	39068	4 AAK71820	Aak71820 Human imm
13	86.4	25.8	39068	4 AAK85294	Aak85294 Human imm
14	86.4	25.8	39068	4 AAK73078	Aak73078 Human imm
15	86.4	25.8	39068	4 AAK87544	Aak87544 Human imm
16	86.4	25.8	39068	7 ABZ68184	Abz68184 Human sec
17	86.4	25.8	39068	7 ABZ74662	Abz74662 Secreted
18	85.8	25.6	45300	4 AAK87547	Aak87547 Human imm
19	85.8	25.6	45300	4 AAK73079	Aak73079 Human imm
20	85.8	25.6	45300	7 ABZ68188	Abz68188 Human sec
21	85.8	25.6	45300	7 ABZ74666	Abz74666 Secreted
22	85.2	25.4	2185	4 AAS32867	Aas32867 Human gen
23	85	25.4	110000	7 AAL52246_2	Continuation (3 of

ALIGNMENTS

RESULT 1

AAA94789
ID AAA94789 standard; DNA; 335 BP.
XX
AC AAA94789;
XX
DT 19-JAN-2001 (first entry)
XX
DE Human Prostate-Specific Membrane antigen Enhancer PSME core region.
XX
KW Core region; human; prostate-specific membrane antigen; enhancer; PSME;
KW cancer; renal cell carcinoma; transitional cell; colonic; neuroendocrine;
KW malignant melanoma; pancreatic duct; breast; soft tissue; PSMA;
KW non-small cell lung; testicular embryonal; glioblastoma multiforme;
KW prostate; breast; bladder; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200052156-A1.
XX
PD 08-SEP-2000.
XX
PF 01-MAR-2000; 2000WO-AU000143.
XX
PR 01-MAR-1999; 99AU-00008956.
PR 25-JAN-2000; 2000AU-00005268.
XX
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX
PI Molloy PL, Watt F;
XX
DR WPI; 2000-594182/56.
XX
PT New regulatory constructs comprising intron 3 of the prostate-specific
PT membrane antigen gene and a heterologous peptide, for directing gene
PT expression in a prostate, bladder, breast or vascular endothelial cell.
XX
PS Claim 10; Fig 11; 56pp; English.
XX
CC The present sequence is the core region from human prostate-specific
CC membrane antigen enhancer (PSME). This sequence is located in intron 3 of
CC the PSMA gene. This sequence was used to construct a recombinant
CC expression cassette, which is useful for the expression of proteins in
CC vascular endothelial cells. In addition, the recombinant construct is
CC useful in the treatment of cancer e.g. renal cell carcinoma, transitional
CC cell carcinoma, colonic adenocarcinoma, neuroendocrine carcinoma,

c	24	85	25.4	237326	2	AAV57903	Aav57903 Hereditar
	25	84.8	25.3	227968	6	ABK83497	Abk83497 Human cDN
	26	84.2	25.1	5690	4	AAK81236	Aak81236 Human imm
	27	84.2	25.1	5690	4	AAK81239	Aak81239 Human imm
	28	84.2	25.1	8724	4	AAK68956	Aak68956 Human imm
	29	83.8	25.0	72604	2	AAZ10752	Aaz10752 Genomic s
c	30	83.8	25.0	121600	6	ABK43231	Abk43231 Human HKN
	31	83.8	25.0	121600	6	ABT10748	Abt10748 Human bre
	32	83.6	25.0	18998	4	AAK36452	Aas36452 Human car
	33	83.6	25.0	18998	9	ADE47146	Ade47146 Human car
	34	83.6	25.0	108316	9	ADC87336	Adc87336 Human GPC
	35	83.2	24.8	28720	2	AAV49655	Aav49655 Human SC3
	36	82.4	24.6	147309	6	ABK49450	Abk49450 Human tra
c	37	82.4	24.6	183610	7	ACF62736	Acf62736 Cancer ba
	38	82.4	24.6	183610	7	ADB20851	Adb20851 MRP1 base
	39	82.4	24.6	183610	9	ADB87940	Adb87940 Human UGT
	40	82.4	24.6	183610	9	ADB96923	Adb96923 Human MDR
	41	82.4	24.6	183610	9	ADB92114	Adb92114 Human MDR
	42	82.2	24.5	21010	4	AAK89247	Aak89247 Human dig
	43	82.2	24.5	21010	4	AAL05888	Aal05888 Human rep
	44	82.2	24.5	21010	4	ABL98452	Ab198452 Human tes
	45	82.2	24.5	21010	7	ABZ68148	Abz68148 Human sec

PR	11-JUL-2000;	2000US-0217487P
PR	11-JUL-2000;	2000US-0217496P
PR	14-JUL-2000;	2000US-0218290P
PR	26-JUL-2000;	2000US-0220963P
PR	26-JUL-2000;	2000US-0220964P
PR	14-AUG-2000;	2000US-0224518P
PR	14-AUG-2000;	2000US-0224519P
PR	14-AUG-2000;	2000US-0225213P
PR	14-AUG-2000;	2000US-0225214P
PR	14-AUG-2000;	2000US-0225266P
PR	14-AUG-2000;	2000US-0225268P
PR	14-AUG-2000;	2000US-0225269P
PR	14-AUG-2000;	2000US-0225270P
PR	14-AUG-2000;	2000US-0225447P
PR	14-AUG-2000;	2000US-0225757P
PR	23-AUG-2000;	2000US-0227182P
PR	23-AUG-2000;	2000US-0227109P
PR	30-AUG-2000;	2000US-0228924P
PR	14-AUG-2000;	2000US-0225759P
PR	14-AUG-2000;	2000US-0226279P
PR	18-AUG-2000;	2000US-0226681P
PR	22-AUG-2000;	2000US-0226686P
PR	22-AUG-2000;	2000US-0227182P
PR	23-AUG-2000;	2000US-0227109P
PR	30-AUG-2000;	2000US-0228924P
PR	01-SEP-2000;	2000US-0229287P
PR	01-SEP-2000;	2000US-0229343P
PR	01-SEP-2000;	2000US-0229344P
PR	01-SEP-2000;	2000US-0229345P
PR	05-SEP-2000;	2000US-0229509P
PR	05-SEP-2000;	2000US-0229513P
PR	06-SEP-2000;	2000US-0230437P
PR	06-SEP-2000;	2000US-0230438P
PR	08-SEP-2000;	2000US-0231242P
PR	08-SEP-2000;	2000US-0231243P
PR	08-SEP-2000;	2000US-0231244P
PR	08-SEP-2000;	2000US-02311413P
PR	08-SEP-2000;	2000US-02311414P
PR	08-SEP-2000;	2000US-0232080P
PR	08-SEP-2000;	2000US-0232081P
PR	12-SEP-2000;	2000US-0231968P
PR	14-SEP-2000;	2000US-0232397P
PR	14-SEP-2000;	2000US-0232398P
PR	14-SEP-2000;	2000US-0232399P
PR	14-SEP-2000;	2000US-0232400P
PR	14-SEP-2000;	2000US-0232401P
PR	14-SEP-2000;	2000US-0233063P
PR	14-SEP-2000;	2000US-0233064P
PR	14-SEP-2000;	2000US-0233065P
PR	14-SEP-2000;	2000US-0233239P
PR	21-SEP-2000;	2000US-0234223P
PR	21-SEP-2000;	2000US-0234274P
PR	25-SEP-2000;	2000US-0234997P
PR	25-SEP-2000;	2000US-0234998P
PR	26-SEP-2000;	2000US-0235484P
PR	27-SEP-2000;	2000US-0235834P
PR	27-SEP-2000;	2000US-0235836P
PR	29-SEP-2000;	2000US-0236327P
PR	29-SEP-2000;	2000US-0236367P
PR	29-SEP-2000;	2000US-0236368P
PR	29-SEP-2000;	2000US-0236369P
PR	29-SEP-2000;	2000US-0236370P
PR	02-OCT-2000;	2000US-0237080P
PR	02-OCT-2000;	2000US-0237037P
PR	02-OCT-2000;	2000US-0237038P
PR	02-OCT-2000;	2000US-0237039P
PR	02-OCT-2000;	2000US-0237040P
PR	13-OCT-2000;	2000US-0239353P
PR	13-OCT-2000;	2000US-0239357P
PR	20-OCT-2000;	2000US-0240960P
PR	20-OCT-2000;	2000US-0241121P
PR	20-OCT-2000;	2000US-0241178P
PR	20-OCT-2000;	2000US-0241786P
PR	20-OCT-2000;	2000US-0241787P
PR	20-OCT-2000;	2000US-0241808P
PR	20-OCT-2000;	2000US-0241809P

20-OCT-2000; 2000US-0241826P.
01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249299P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254037P.
05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451937/48.

Isolated polypeptide for treating, preventing and/or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis.

Example 2; SEQ ID NO 3660; 781bp + Sequence Listing; English.

The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,

CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

Sequence 13444 BP; 4456 A; 2639 C; 2226 G; 4123 T; 0 U; 0 Other;

Query Match 28.1%; Score 94; DB 4; Length 13444;

Best Local Similarity 65.0%; Pred. No. 5.4e-09;

Matches 139; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 122 CATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTAT 181
DB 4821 CATCTTAAATTTTCTCTTTTAAATATGATTTCTATTTATTTGGTCTTCAT 4762
QY 182 GTCATAGAACAAATTTTATTTGTTGATATATGACTTTATTTGTTGCTATACAACTA 241
DB 4761 GTTAAATACAGAAATCTGTTTCTAGGTAGATTTTTCAGTGTCAATGGGATTTTCACTTT 4702
QY 242 GATTTTGTGTTTGTGACCGAGTCTTACTCTGTGACCAGGCTGGAGTGTATGCA 301
DB 4701 TTTTGTGTTTGTGACCGAGTCTGCTGTGCGCCAGCGCGAGTGTGCGGCG 4642
QY 302 TGGTCTCAGTCTCACTGCAACTCGCTCCGCTCCGGG 335
DB 4641 TGATCTCGGCTCACTGCAAGCTCGGCTTCGGG 4608

RESULT 3

ABX60283/c

ID ABX60283 standard; cDNA; 13444 BP.

XX AC ABX60283;

XX DT 26-FEB-2003 (first entry)

XX DE cDNA encoding novel human musculoskeletal system antigen #2627.

XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
KW re-vascularisation; thrombosis; arteriosclerosis; mineral content;
KW cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
KW post-operative tissue repair; limb regeneration; neuronal growth;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW AIDS-related complex; chondrocyte growth; bone regeneration;
KW periodontal regeneration; tissue transport; bone graft; skin aging;
KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
KW cell growth; organ transplant; cell differentiation; body height; weight;
KW hair colour; eye colour; skin; percentage of adipose tissue;
KW pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
KW depression; tendency for violence; pain; reproductive capability;
KW hormone level; endocrine level; appetite; libido; memory; stress;
KW storage capability; fat content; lipid content; protein content;
KW carbohydrate content; vitamin content; cofactor content;
KW nutritional component.

XX OS Homo sapiens.

XX PN US2002147140-A1.

XX PD 10-OCT-2002.

XX PF 17-JAN-2001; 2001US-00764877.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 28-JUN-2000; 2000US-0214866P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2000US-0254097P.
PR 05-JAN-2001; 2000US-02559678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis. Disclosure; SEQ ID NO 42366; 307lpp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AA82170 to AA91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome


```
DT 04-DEC-2002 (first entry)
XX
DE Human breast cancer associated coding sequence SEQ ID NO: 280.
XX
KW Human; breast specific gene; breast cancer; differential expression;
KW cytostatic; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200259271-A2.
XX
PD 01-AUG-2002.
XX
PF 25-JAN-2002; 2002WO-US0021176.
XX
PR 25-JAN-2001; 2001US-0263757P.
PR 25-APR-2001; 2001US-0286090P.
PR 23-MAY-2001; 2001US-0292517P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Orr MS, Nation M, Diggans JC, Zeng W;
XX
DR WPI; 2002-674803/72.
XX
PT Diagnosing breast cancer in a patient comprises detecting the level of
PT gene expression in cell or tissue samples, where a differential gene
PT expression is indicative of breast cancer.
XX
PS Claim 1; SEQ ID NO 280; 260pp + Sequence Listing; English.
XX
CC The present invention relates to methods of diagnosing breast cancer in a
CC patient, which comprise detecting the level of expression in a tissue
CC sample of two or more genes selected from those shown in ABT09867-
CC ABT1112, where a differential expression of the genes indicates breast
CC cancer. The methods are useful in diagnosing, treating, detecting the
CC progression, and in monitoring treatment of breast cancer in patients.
CC The methods are also useful as a screening tool for agents that modulate
CC the onset or progression of breast cancer. The breast cancer genes may be
CC used as diagnostic markers for the prediction or identification of the
CC malignant state of breast tissue, for confirming the type and progression
CC of cancer, and for drug screening and assays. The present sequence is a
CC coding sequence of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub.published_pct_sequences
XX
SQ Sequence 140167 BP; 45038 A; 27502 C; 26800 G; 40827 T; 0 U; 0 Other;

Query Match 26.0%; Score 87.2; DB 6; Length 140167;
Best Local Similarity 69.2%; Pred. No. 1.1e-07;
Matches 119; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 164 GTCATATGATATTTATGTATAGAACAAATTTATGCTTGATATATGACTTTATGT 223
Db 123721 GCCAGAGATAAATGCCAGATAGGTAGGTATGTTTGGTATATACATATATATATA 123662

QY 224 TATATGGCTATACACTAGATTTTTTGTGTTTTGACCGAGTCTTACTCTGTCAACCA 283
Db 123661 TATATATATATATATATATGATATTTTTTTTTCAGATGGAGTCTTGTCTGTCAACCA 123602

QY 284 GGCTGGAGTGTATGGCATGGTCTCAGTCACTGCAACCTCGCCTCCCGGG 335
Db 123601 GGCTGGAGTTCAGTGGCATGATCTCAGTCACTGCAAGCCGCCACCGGG 123550

RESULT 12
AAK71820/c
ID AAK71820 standard; DNA; 39068 BP.
XX
AC AAK71820;
XX
DT 06-NOV-2001 (first entry)
```

```
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26632.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
```

```
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236376P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-023935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 26632; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 39068 BP; 10474 A; 9094 C; 8680 G; 10820 T; 0 U; 0 Other;

Query Match 25.8%; Score 86.4; DB 4; Length 39068;
Best Local Similarity 56.8%; Pred. No. 1.5e-07;
Matches 159; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 56 TGTTTGGGAAATTTTAAACTACTGTTTAAATTTCTTTTATTTGTTGTTGTAATGACTATT 115
DB 32919 TATATATACATATATATACATATATATATATATATATATATATATATATATATACA 32860

QY 116 TTACGTCATATAACAATTTTATTTGTTGTTGTTAAATGACTTTTATTTGTTGCATATGATAA 175
DB 32859 TATATACATATATACATATATATATATATATATATATATATATATATATATATACATA 32800

QY 176 TTTTATGTCATAGAACAAATTTTATTTGTTGTTGATATATGACTTTTATTTATATGCTTATA 235
DB 32799 TACATACATATACATATATATATATATATATATATATATATATATATATATATATGTA 32740

QY 236 CAACTAGATTTTGTGTTGTTTGTACCGAGTCTTACTCTGTCCACCGAGCTGGAGTGTA 295
DB 32739 TATATATTTTGTGTTGTTTGTAGACAGTGCACCTCTGTCCACCGAGCTGGAGTGCA 32680

QY 296 ATGGCATGGTCTCAGCTCACTGCAACCTCCGCTCCCGGG 335
DB 32679 GTGGCAGAGTCTCAGCTCACTGCGAGTCTCTGCTCCCGAGG 32640

RESULT 13
AAK85294/c
ID AAK85294 standard; DNA; 39068 BP.
XX
XX AAK85294;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40106.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
```

XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001354.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0235447P.
PR 14-AUG-2000; 2000US-0235757P.
PR 14-AUG-2000; 2000US-0235758P.
PR 14-AUG-2000; 2000US-0235759P.
PR 18-AUG-2000; 2000US-0236279P.
PR 22-AUG-2000; 2000US-0236681P.
PR 22-AUG-2000; 2000US-0236868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0234999P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-02559678P.
XX

PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226581P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231142P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231144P.
PR 08-SEP-2000; 2000US-0231145P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234397P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX XX

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and metastasis.

Disclosure; SEQ ID NO 42356; 3071pp + Sequence Listing; English.

[illegible]

Search completed: May 21, 2004, 17:36:52
Job time : 142.729 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 16:26:10 ; Search time 1344.31 Seconds
(without alignments)
10801.019 Million cell updates/sec

Title: US-09-914-651A-1

Perfect score: 335
Sequence: 1 aattattttcttctaacc.....tgcaactccgcctccggg 335

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.lnv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	335	100.0	335	6	BD269371	BD269371 Controlli
2	335	100.0	93525	9	AF007544	AF007544 Homo sapi
3	335	100.0	117521	2	AC023784	AC023784 Homo sapi
4	335	100.0	129095	9	AC110742	AC110742 Homo sapi
5	335	100.0	158524	2	AL162372	AL162372 Homo sapi
C 6	335	100.0	187529	9	AC118273	AC118273 Homo sapi
7	327	97.6	1669	9	AF480875	AF480875 Homo sapi
8	327	97.6	135637	9	AP003400	AP003400 Homo sapi
9	327	97.6	137888	9	AP003122	AP003122 Homo sapi
C 10	327	97.6	192848	2	AC024234	AC024234 Homo sapi
11	222.8	66.5	246865	2	AC074003	AC074003 Homo sapi
12	100	29.9	161940	2	AC079289	AC079289 Homo sapi
13	99	29.6	125685	9	AC073310	AC073310 Homo sapi
C 14	95.8	28.6	184864	9	AC013553	AC013553 Homo sapi
15	95.6	28.5	163401	2	AC073251	AC073251 Homo sapi
16	95.6	28.5	167878	9	AC096645	AC096645 Homo sapi
17	95.2	28.4	91792	9	AC092355	AC092355 Homo sapi
C 18	94.8	28.3	161925	9	AC090103	AC090103 Homo sapi
19	94.6	28.2	103930	9	AL358234	AL358234 Human DNA
20	94.6	28.2	113075	9	AL512363	AL512363 Human DNA
C 21	94.6	28.2	190202	9	AC022021	AC022021 Homo sapi
22	94.6	28.2	192952	2	AC028966	AC028966 Homo sapi
C 23	94	28.1	113729	9	AF260011	AF260011 Homo sapi
24	94	28.1	166977	9	AC007198	AC007198 Homo sapi
C 25	94	28.1	340000	9	HS21C048	AL163248 Homo sapi
C 26	93.8	28.0	198935	2	AC068573	AC068573 Homo sapi
27	93.4	27.9	106954	9	AC104651	AC104651 Homo sapi
C 28	93.2	27.8	217985	9	AC021193	AC021193 Homo sapi
29	93	27.8	181413	2	AP001352	AP001352 Homo sapi
30	92.8	27.7	154639	2	AC090175	AC090175 Homo sapi
31	92.6	27.6	107415	9	AL513365	AL513365 Human DNA
C 32	92.4	27.5	169240	9	AC022069	AC022069 Homo sapi
33	92	27.5	103566	9	AC097481	AC097481 Homo sapi
34	91.6	27.3	87791	9	HSJ223D17	AL121882 Human DNA
35	91.6	27.3	167869	2	AC021339	AC021339 Homo sapi
36	91.6	27.3	213044	9	AC096576	AC096576 Homo sapi
C 37	91.6	27.3	240379	9	AC005077	AC005077 Homo sapi
38	91.4	27.3	143770	9	AL589987	AL589987 Human DNA
39	91.2	27.2	102910	9	AB041512	AB041512 Homo sapi
C 40	91.2	27.2	144234	2	AP002425	AP002425 Homo sapi
41	91.2	27.2	144279	2	AP001570	AP001570 Homo sapi
42	91.2	27.2	162217	9	AC068043	AC068043 Homo sapi
43	91.2	27.2	216972	9	AC067819	AC067819 Homo sapi
44	91	27.2	162500	9	AC104513	AC104513 Homo sapi
C 45	91	27.2	167434	9	AC079910	AC079910 Homo sapi

ALIGNMENTS

RESULT 1
BD269371
LOCUS
DEFINITION
BD269371 Controlling construct containing intron 3 of prostate-specific
membrane antigen gene.
ACCESSION
BD269371
VERSION
BD269371.1 GI:33079139
KEYWORDS
JP 2002537807-A/1.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 335)
Molloy,P.L. and Watt,F.
AUTHORS
Controlling construct containing intron 3 of prostate-specific

BD269371 335 bp DNA linear PAT 17-JUL-2003
Controlling construct containing intron 3 of prostate-specific

```
JOURNAL      membrane antigen gene
COMMENT      Patent: JP 2002537807-A 1 12-NOV-2002;
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION
OS           Homo sapiens (human)
PN           JP 2002537807-A/1
PD           12-NOV-2002
PF           01-MAR-2000 JP 2000502768
PR           01-MAR-1999 AU PP 8956,25-JAN-2000 AU PQ 5268 PI
PETER LAURENCE MOLLOY,FUJIKO WATT
PC           C12N15/09,A61K38/00,A61K38/46,A61K48/00,A61P35/00,C12N5/10, PC
C12N15/00,
PC           A61K37/02,A61K37/54,C12N5/00
CC           Controlling construct containing intron
CC           3 of prostate-specific
CC           membrane
CC           antigen gene
CC           Location/Qualifiers
FT           key 1..335
FT           source /organism='Homo sapiens (human)'.
FEATURES
  source      Location/Qualifiers
    1..335
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 335; DB 6; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.1e-48;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATTATTTTTCCTTTAACTCTTCAAACTCAAGAAACAGTGGCTTGACTCTGTTT 60
DB 1 AATTATTTTTCCTTTAACTCTTCAAACTCAAGAAACAGTGGCTTGACTCTGTTT 60
QY 61 GTGAAAATTTTAACTACTGTTTAAATTTCTTTTATTGTTGTAATGACTATTTTACG 120
DB 61 GTGAAAATTTTAACTACTGTTTAAATTTCTTTTATTGTTGTAATGACTATTTTACG 120
QY 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTTGTTGTCATATGATAATTTTA 180
DB 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTTGTTGTCATATGATAATTTTA 180
QY 181 TGTCATAGACAATTTTATTGTTGTAATGACTTTATTTGTTGTAATGACTATTTTACG 240
DB 181 TGTCATAGACAATTTTATTGTTGTAATGACTTTATTTGTTGTAATGACTATTTTACG 240
QY 241 AGATTTTGTGTTTGTGTTTGTGACCGAGTCTTACTCTGTCCACCGCTGGAAGTGTAAATGGC 300
DB 241 AGATTTTGTGTTTGTGTTTGTGACCGAGTCTTACTCTGTCCACCGCTGGAAGTGTAAATGGC 300
QY 301 ATGGTCTCAGCTCACTGCAACCTCCGCTCCCGG 335
DB 301 ATGGTCTCAGCTCACTGCAACCTCCGCTCCCGG 335
RESULT 2
AF007544
LOCUS      Homo sapiens prostate-specific membrane antigen (PSM) gene,
DEFINITION complete cds.
ACCESSION AF007544
VERSION   AF007544.1 GI:2970122
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 93525)
AUTHORS   O'Keefe,D.S., Su,S.L., Luo,Y., Horiguchi,Y., Bacich,D.J., Luo,Y.,
Powell,C.T., Zandvliet,D., Russell,P.J., Molloy,P.L., Nowak,N.J.,
Shows,T.B., Mullins,C., Vonder Haar,R.A., Fair,W.R. and Heston,W.D.
TITLE     Mapping, genomic organization and promoter analysis of the human
prostate-specific membrane antigen gene
Biochim. Biophys. Acta 1443 (1-2), 113-127 (1998)
99057588
2 (bases 1682 to 65315)
Heston,W.D.W., Su,S.L., Luo,Y., Huryk,R., Bacich,D.J., Fair,W.R.,
Mullins,C. and Vonder Haar,R.A.
Direct Submission
Submitted (06-JUN-1997) Memorial Sloan Kettering Cancer Center,
1275 York Avenue, New York, New York 10021, USA
3 (bases 1 to 93525)
O'Keefe,D.S., Su,S.L., Luo,Y., Horiguchi,Y., Bacich,D.J.,
Powell,C.T., Zandvliet,D., Russell,P.J., Molloy,P.L., Nowak,N.J.,
Mullins,C., Vonder Haar,R.A., Fair,W.R. and Heston,W.D.W.
Direct Submission
Submitted (18-MAR-1998) Memorial Sloan Kettering Cancer Center,
1275 York Avenue, New York, New York 10021, USA
REMARK     Sequence update by submitter
COMMENT     On Mar 19, 1998 this sequence version replaced gi:2897945.
FEATURES
  Location/Qualifiers
    1..93525
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11p11-12"
    2488..64518
    /genes="PSM"
    join(2488..2863,4994..5099,10726..10912,18275..18376,
    24400..24525,25314..25500,27927..28020,35216..35314,
    36196..36281,37697..37816,39896..39978,41911..41974,
    4402..46469,53129..53220,54364..54454,56661..56925,
    57226..57307,62423..62515,64209..64398)
    /genes="PSM"
    /product="prostate-specific membrane antigen"
    join(2745..2863,4994..5099,10726..10912,18275..18376,
    24400..24525,25314..25500,27927..28020,35216..35314,
    36196..36281,37697..37816,39896..39978,41911..41974,
    4402..46469,53129..53220,54364..54454,56661..56925,
    57226..57307,62423..62515,64209..64398)
    /genes="PSM"
    /note="PSMA"
    /codon_start=1
    /product="prostate-specific membrane antigen"
    /protein_id="AAC83972.1"
    /db_xref="GI:2897946"
    /translation="MNLLHETDSAVATARRPRWLCAALVLGGFFLLQFLRGWPK
    SNEATNTPKNNKAFDLKAEINIKPLHNTQIPLHAGTEQNFQAKIQOSQWKE
    FGLDSVELAHYDVLISYPNKTNPYISINEDGNEIFNTSLPEPPGKYENVSIDVPP
    LAGAGVILSDPADYFAPGVKSYDGNLPGGVQVQGNILNGLAGDPLTPGPANE
    YARRGIAEAVGLSPVHPHIGYDAQKLEKMGSGAPPDSSWRGSLKVPYVPGFT
    GNFTOKVWHIHSNTEVTRIYNVIGTLRGVAPDRYVILGHRDSWFGGIDPQGA
    AVVHEIVRSFGTLKKEGWRPRTILPASMDAREFGLLGSTEAENSRLLQRRGAVI
    NADSSLEGNYTLRVDCITPLMYSLVHNLTKELSPDEGFGKSLYESWTKSPSPFSG
    MPRIKLGSGNDVFVFRQIGIASGRARYTKWETNKFSGYPLDIHYSVETVELKPY
    DPMFKYHLTVAQVQGMVPELANSIVLPFCRDYAVVLKADIKYISIMKHQEMKT
    YSVSDFLSAVKNFETIASKFSRQDQKSNPIVLRMNDQMLEAFIDPLGIP
    DRPFYRHVIYAPSSHNVKAGSPGGIYDALFDIESKVPDSKAWGEVKRQIYVAFTVQ
    AAATLSEVA"
ORIGIN
Query Match 100.0%; Score 335; DB 9; Length 93525;
Best Local Similarity 100.0%; Pred. No. 3.2e-49;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATTATTTTTCCTTTAACTCTTCAAACTCAAGAAACAGTGGCTTGACTCTGTTT 60
DB 14760 AATTATTTTTCCTTTAACTCTTCAAACTCAAGAAACAGTGGCTTGACTCTGTTT 14819
QY 61 GTGAAAATTTTAACTACTGTTTAAATTTCTTTTATTGTTGTAATGACTATTTTACG 120
DB 14820 GTGAAAATTTTAACTACTGTTTAAATTTCTTTTATTGTTGTAATGACTATTTTACG 14879
```



```
repeat_region /rpt_family="L1ME4A"
10670..10691
repeat_region /rpt_family="AT_rich"
10977..10997
repeat_region /rpt_family="AT_rich"
complement(10998..11293)
repeat_region /rpt_family="AluJo"
12004..12042
repeat_region /rpt_family="AT_rich"
12337..12396
repeat_region /rpt_family="(TA)n"
12447..12489
repeat_region /rpt_family="AT_rich"
12950..12970
repeat_region /rpt_family="AT_rich"
complement(13047..13097)
repeat_region /rpt_family="MIR3"
13218..13339
repeat_region /rpt_family="MIR"
13460..13500
repeat_region /rpt_family="AT_rich"
14066..14087
repeat_region /rpt_family="AT_rich"
14141..14166
repeat_region /rpt_family="(TTTTTG)n"
15042..15277
repeat_region /rpt_family="AluSp"
complement(16069..16260)
repeat_region /rpt_family="MIR"
complement(16362..16449)
repeat_region /rpt_family="MIR"
16603..16637
repeat_region /rpt_family="(TG)n"
16976..17065
repeat_region /rpt_family="L1"
17264..17286
repeat_region /rpt_family="(A)n"
17817..17860
repeat_region /rpt_family="(TTTG)n"
18317..18432
repeat_region /rpt_family="L2"
18922..19089
repeat_region /rpt_family="MER45A"
20139..20250
repeat_region /rpt_family="FLAM_C"
20252..20292
repeat_region /rpt_family="(TAAAA)n"
20804..21817
repeat_region /rpt_family="L1MA3"
21815..22853
repeat_region /rpt_family="L1MA3"
22833..24145
repeat_region /rpt_family="L1MA3"
24826..24896
repeat_region /rpt_family="AT_rich"
25215..25380
repeat_region /rpt_family="MIR"
25381..25402
repeat_region /rpt_family="(TG)n"
25403..25440
repeat_region /rpt_family="MIR"
complement(26317..26407)

Query Match 100.0%; Score 335; DB 9; Length 129095;
Best Local Similarity 100.0%; Pred. No. 2.9e-49;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATATTTTTCCTTAACCTTTCAAACTCAAGGAAACAGTTGGCCCTTGACTCTGTTT 60
Db 123370 AATATTTTTCCTTAACCTTTCAAACTCAAGGAAACAGTTGGCCCTTGACTCTGTTT 123429
QY 61 GTGGAATTTTAACTACTGTTTAAATTTCTTTATGTTGTAATGACTATTTTACG 120
|||||
```

```
Db 123430 GTGGAATTTTAACTACTGTTTAAATTTCTTTATGTTGTAATGACTATTTTACG 123489
QY 121 TCATATAACAATTTTATTTTATTTGTTTAAATGACTTTATTTGTTTCATATGATAATTTTA 180
|||||
Db 123490 TCATATAACAATTTTATTTTATTTGTTTAAATGACTTTATTTGTTTCATATGATAATTTTA 123549
|||||
QY 181 TGTATAGAACAAATTTTATTTGTTTATGATATATGACTTTATTTGTTTATGATGCTATCAACT 240
|||||
Db 123550 TGTATAGAACAAATTTTATTTGTTTATGATATATGACTTTATTTGTTTATGATGCTATCAACT 123609
|||||
QY 241 AGATTTTTTTGTTGTTTGGACCGAGTCTTACTCTGTCAACCCAGGCTGGAGTGAATGCG 300
|||||
Db 123610 AGATTTTTTTGTTGTTTGGACCGAGTCTTACTCTGTCAACCCAGGCTGGAGTGAATGCG 123669
|||||
QY 301 ATGCTCTCAGCTCACTGCAACTCGGCTCGCGGG 335
|||||
Db 123670 ATGCTCTCAGCTCACTGCAACTCGGCTCGCGGG 123704
|||||

RESULT 5
AL162372
LOCUS Homo sapiens chromosome 13 clone RP11-181P16, 3 unordered pieces.
DEFINITION AL162372
ACCESSION AL162372.11 GI:12539555
VERSION HTG; HTGS PHASE1; HTGS_CANCELLED.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Burton, J.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jan 25, 2001 this sequence version replaced gi:11693478.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bA181P16
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: pGAP4; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 157639 bases at least Q40
Consensus quality: 157977 bases at least Q30
Consensus quality: 158153 bases at least Q20
Insert size: 158324; sum-of-contigs
Insert size: 174362; 5.7% error; agarose-fp
Quality coverage: 6.90x in Q20 bases; sum-of-contigs Quality
coverage: 6.26x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* As soon as it is available and the accession number will
* be preserved.
* 1 59356: contig of 59356 bp in length
* 59357 59456: gap of 100 bp
* 59457 65248: contig of 5792 bp in length
* 65249 65348: gap of 100 bp
* 65349 158524: contig of 93176 bp in length.
* Location/Qualifiers
* 1..158524
* /organism="Homo sapiens"
* /mol_type="genomic DNA"

FEATURES
source
```

/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-181P16"
/clone_lib="RP11-11.1"
1. 59356
/note="assembly_fragment:00306
clone_end:T7
vector_side:left"
59457_-65248
/note="assembly_fragment:01282
fragment_chain:1"
65349_-158524
/note="assembly_fragment:02815
fragment_chain:1"

ORIGIN

Query Match 100.0%; Score 335; DB 2; Length 158524;
Best Local Similarity 100.0%; Pred. No. 2.8e-49;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTATTTTTCCTTTAACTTTCAAACTCAAGGAAAACAGTTGGCTTGACTCTGTTT 60
DB 71493 AATTATTTTTCCTTTAACTTTCAAACTCAAGGAAAACAGTTGGCTTGACTCTGTTT 71552

QY 61 GTGAAAATTTTAACTACTGTTTAAATTTCTTTATTTGTTGTAATGACTATTTTACG 120
DB 71553 GTGAAAATTTTAACTACTGTTTAAATTTCTTTATTTGTTGTAATGACTATTTTACG 71612

QY 121 TCATATAACAATTTTATTTGTTTAAAGACTTTATTTGTTTGTATATGATAAATTTTA 180
DB 71613 TCATATAACAATTTTATTTGTTTAAAGACTTTATTTGTTTGTATATGATAAATTTTA 71672

QY 181 TGCATAGAACAAATTTTATTTGTTGATATATGACTTTATTTGTTATATGCTATACAACT 240
DB 71673 TGCATAGAACAAATTTTATTTGTTGATATATGACTTTATTTGTTATATGCTATACACT 71732

QY 241 AGATTTTGTGTTTTCACCGAGCTTACTCTGTACCCAGGCTGGAGTGTATAGGC 300
DB 71733 AGATTTTGTGTTTTCACCGAGCTTACTCTGTACCCAGGCTGGAGTGTATAGGC 71792

QY 301 ATGGTCTCACTCACTCACTCCGCTCCCGG 335
DB 71793 ATGGTCTCACTCACTCACTCCGCTCCCGG 71827

RESULT 6
AC118273/c
LOCUS
DEFINITION Homo sapiens chromosome 11, clone RP11-107P7, complete sequence.
ACCESSION AC118273
VERSION AC118273.4 GI:22759265
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 187529)
AUTHORS Birren,B., Nussbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-107P7
JOURNAL Unpublished
COMMENT Submitted (09-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2002 this sequence version replaced gi:22539138.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26235
Center clone name: 107_P_7
----- Location/Qualifiers
1. 187529
/organism="Homo sapiens"

Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 187529)
Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted (29-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 187529)
Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted (09-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2002 this sequence version replaced gi:22539138.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26235
Center clone name: 107_P_7
----- Location/Qualifiers
1. 187529
/organism="Homo sapiens"

versus a Prostate-Specific Membrane Antigen Like Gene in Normal
Tissues, Prostate Cancer and Tumor Associated-Vasculature
Unpublished

REFERENCE 2 (bases 1 to 1669)

AUTHORS O'Keefe,D.S., Bacich,M.E., Bacich,D.J. and Heston,W.D.W.

TITLE Direct Submission

JOURNAL Submitted (05-FEB-2002) Cancer Biology, Cleveland Clinic
Foundation, Box ND50, 9500 Euclid Ave., Cleveland, OH 44195, USA

FEATURES Location/Qualifiers

source

1..1669

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q14.3"

misc_feature

1..1669

/note="prostate-specific membrane antigen-like enhancer

region; PSMA-like; FOLHI-like"

ORIGIN

Query Match 97.6%; Score 327; DB 9; Length 1669;
Best Local Similarity 98.5%; Pred. No. 1.9e-47;
Matches 330; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AATTATTTTTCCTTAACTTCAAACTCAAGAAACCAGTTGGCTTGACTCTCTTT 60

Db 58 AATTATTTTTCCTTAACTTCAAACTCAAGAAACCAGTTGGCTTGACTCTCTTT 117

QY 61 GTGAAAATTTTAACTACTGGTTAAATTCCTTTATTTGGTTGTAATGACTATTTTACG 120

Db 118 GTGAAAATTTTAACTACTGGTTAAATTCCTTTATTTGGTTGTAATGACTATTTTATG 177

QY 121 TCATATAACAATTTTATTTGTTGTAATGACTTTTATTTGTTGTCATATGATAAATTTTA 180

Db 178 TCATATAACAATTTTATTTGTTGTAATGACTTTTATTTGTTGTCATATGATAAATTTTA 237

QY 181 TGTCTATAGACAATTTTATTTGTTGTAATGACTTTTATTTGTTGTAATGACTTTTATG 240

Db 238 TGTCTATAGACAATTTTATTTGTTGTAATGACTTTTATTTGTTGTAATGACTTTTATG 297

QY 241 AGATTTTGTGTTTTCACCGAGTCTTACTCTGTCCACCGCTGGAGTGAATGGC 300

Db 298 AGATTTTGTGTTTTCACCGAGTCTTACTCTGTCCACCGCTGGAGTGAATGGC 357

QY 301 ATGTCTCTCAGTCACTCAACCTCCGCTCCCGG 335

Db 358 ATGTCTCTCAGTCACTCAACCTCCGCTCCCGG 392

RESULT 8 135637 bp DNA linear PRI 15-MAR-2003

AP003400 Homo sapiens genomic DNA, chromosome 11q, clone:RP11-643G5,

LOCUS complete sequence.

DEFINITION AP003400

VERSION AP003400.2 GI:15320508

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Homo sapiens genomic DNA

JOURNAL Published Only in Database (2001)

REFERENCE 2 (bases 1 to 135637)

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Direct Submission

JOURNAL Submitted (08-MAR-2001) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)

On Aug 27, 2001 this sequence version replaced gi:13366104.

COMMENT

FEATURES

source

1..135637

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q"

/clone="RP11-643G5"

ORIGIN

Query Match 97.6%; Score 327; DB 9; Length 135637;
Best Local Similarity 98.5%; Pred. No. 7e-48;
Matches 330; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AATTATTTTTCCTTAACTTCAAACTCAAGAAACCAGTTGGCTTGACTCTCTTT 60

Db 133621 AATTATTTTTCCTTAACTTCAAACTCAAGAAACCAGTTGGCTTGACTCTCTTT 133680

QY 61 GTGAAAATTTTAACTACTGGTTAAATTCCTTTATTTGGTTGTAATGACTATTTTACG 120

Db 133681 GTGAAAATTTTAACTACTGGTTAAATTCCTTTATTTGGTTGTAATGACTATTTTATG 133740

QY 121 TCATATAACAATTTTATTTGTTGTAATGACTTTTATTTGTTGTCATATGATAAATTTTA 180

Db 133741 TCATATAACAATTTTATTTGTTGTAATGACTTTTATTTGTTGTCATATGATAAATTTTA 133800

QY 181 TGTCTATAGACAATTTTATTTGTTGTAATGACTTTTATTTGTTGTAATGACTTTTATG 240

Db 133801 TGTCTATAGACAATTTTATTTGTTGTAATGACTTTTATTTGTTGTAATGACTTTTATG 240

QY 241 AGATTTTGTGTTTTCACCGAGTCTTACTCTGTCCACCGCTGGAGTGAATGGC 300

Db 133861 AGATTTTGTGTTTTCACCGAGTCTTACTCTGTCCACCGCTGGAGTGAATGGC 133920

QY 301 ATGTCTCTCAGTCACTCAACCTCCGCTCCCGG 335

Db 133921 ATGTCTCTCAGTCACTCAACCTCCGCTCCCGG 133955

RESULT 9 137888 bp DNA linear PRI 15-MAR-2003

AP003122 Homo sapiens genomic DNA, chromosome 11q, clone:RP11-31312,

LOCUS complete sequence.

DEFINITION AP003122

VERSION AP003122.2 GI:15320502

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Homo sapiens genomic DNA

JOURNAL Published Only in Database (2001)

REFERENCE 2 (bases 1 to 137888)

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Direct Submission

JOURNAL Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,

Tel:81-45-503-9111, Fax:81-45-503-9170)

On Aug 27, 2001 this sequence version replaced gi:12597178.

COMMENT

FEATURES

source

1..137888

/organism="Homo sapiens"

/mol_type="genomic DNA"

```
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-31312"

ORIGIN
Query Match          97.6%; Score 327; DB 9; Length 137888;
Best Local Similarity 98.5%; Pred. No. 7e-48;
Matches 330; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AATTATTTTTCCTTTAACTTTTAACTTTCAAGGAAACCAAGTTCGCTTACTCTGTTT 60
DB 825 AATTATTTTTCCTTTAACTTTTAACTTTCAAGGAAACCAAGTTCGCTTACTCTGTTT 884

QY 61 GTGGAATTTTAACTACTGTTTAACTTTTAACTTTTAACTTTTAACTACTGTTTAACT 120
DB 885 GTGGAATTTTAACTACTGTTTAACTTTTAACTTTTAACTTTTAACTACTGTTTAACT 944

QY 121 TCATATACAAATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 180
DB 945 TCATATACAAATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 1004

QY 181 TGTATAGAACAAATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 240
DB 1005 TGTATAGAACAAATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 1064

QY 241 AGATTTTTCGTTGTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 300
DB 1065 AGATTTTTCGTTGTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 1124

QY 301 ATGCTCTCAGCTCACTGCAACTTCGCTTCGCGG 335
DB 1125 ATGCTCTCAGCTCACTGCAACTTCGCTTCGCGG 1159

RESULT 10
AC024234/5 192648 bp DNA linear HTG 01-SEP-2000
LOCUS Homo sapiens chromosome 11 clone RP11-31312, WORKING DRAFT
DEFINITION AC024234
ACCESSION AC024234
VERSION AC024234.5 GI:8569915
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 192648)
Waterston,R.H.
Direct Submission
Submitted (28-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jun 16, 2000 this sequence version replaced gi:7284696.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0311302
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing method: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 181493 bases at least Q40
Consensus quality: 184864 bases at least Q20

----- Location/Qualifiers -----
1..192648
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"

Consensus quality: 186864 bases at least Q20
Insert size: 18000; agarose-fp
Quality coverage: 3.74 in Q20 bases; agarose-fp
Quality coverage: 3.73 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1413: contig of 1413 bp in length
1513: gap of unknown length
1514 1978: contig of 1465 bp in length
2979 3078: gap of unknown length
3079 4859: contig of 1781 bp in length
4860 6723: contig of unknown length
6724 9068: contig of 2245 bp in length
9069 9169: gap of unknown length
9170 11510: contig of 2342 bp in length
11511 15431: contig of 3821 bp in length
15432 18469: contig of unknown length
18470 22092: contig of 3523 bp in length
22093 25250: contig of 3058 bp in length
25251 28352: contig of 3002 bp in length
28353 31550: contig of 3097 bp in length
31551 34632: contig of 2982 bp in length
34633 37371: gap of unknown length
37372 38568: gap of unknown length
38569 42337: contig of 3769 bp in length
42338 47543: contig of 5106 bp in length
47544 56048: contig of 8405 bp in length
56049 64671: contig of 8523 bp in length
64672 73100: contig of 8329 bp in length
73101 81814: contig of 8614 bp in length
81815 91133: contig of 9219 bp in length
91134 101427: contig of 10194 bp in length
101428 101527: gap of unknown length
101528 113041: contig of 11514 bp in length
113042 125504: contig of 12363 bp in length
125505 140510: contig of 14906 bp in length
140511 140611: gap of unknown length
140612 158205: contig of 17594 bp in length
158206 192648: contig of 34344 bp in length.
Location/Qualifiers
1..192648
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
```

```
/clone="RP11-31312"
1. .1413
/note="assembly_name:Contig10"
1514. .2978
/note="assembly_name:Contig11"
3079. .4859
/note="assembly_name:Contig12"
4960. .6723
/note="assembly_name:Contig13"
6824. .9068
/note="assembly_name:Contig14"
9169. .11510
/note="assembly_name:Contig15"
11611. .15431
/note="assembly_name:Contig16"
15532. .18469
/note="assembly_name:Contig17"
18570. .22092
/note="assembly_name:Contig18"
22193. .25250
/note="assembly_name:Contig19"
25351. .28352
/note="assembly_name:Contig20"
28453. .31549
/note="assembly_name:Contig21"
31650. .34631
/note="assembly_name:Contig22"
34732. .38468
/note="assembly_name:Contig23"
38569. .42337
/note="assembly_name:Contig24"
42438. .47543
/note="assembly_name:Contig25"
47644. .56048
/note="assembly_name:Contig26"
56149. .84671
/note="assembly_name:Contig27"
64772. .73100
/note="assembly_name:Contig28"
73201. .81814
/note="assembly_name:Contig29"
81915. .91133
/note="assembly_name:Contig30"
91234. .101427
/note="assembly_name:Contig31"
clone end:SP6
vector side:left"
101526. .113041
/note="assembly_name:Contig32"
113142. .125504
/note="assembly_name:Contig33"
125605. .140510
/note="assembly_name:Contig34"
140611. .158204
/note="assembly_name:Contig35"
158305. .192648
/note="assembly_name:Contig36"

ORIGIN
Query Match 97.6%; Score 327; DB 2; Length 192648;
Best Local Similarity 98.5%; Pred. No. 6.5e-48;
Matches 330; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AATTATTTTTCCTTTAACTCAAACTCAAGGAAACACGATGGCTTGACTCTCTTT 60
Db 174244 AATTATTTTTCCTTTAACTCAAACTCAAGGAAACACGATGGCTTGACTCTCTTT 174185

QY 61 GTGGAATTTTAACTACTGGTTTAACTTTTATTTGTTGTAATGACTATTTTACG 120
Db 174184 GTGGAATTTTAACTACTGGTTTAACTTTTATTTGTTGTTATATGACTATTTTATG 174125

QY 121 TCATATACAAATTTTATTTGTTTAACTGTTTATTTGTTCTCATATGATAATTTTA 180
Db 174124 TCATATACAAATTTTATTTGTTTAACTGTTTATTTGTTCTCATATGATAATTTTA 174065

QY 181 TGTCTATAGAACAAATTTTATTTGTTCTGATATATGACTTTTATTTGTTATATGCTATACAACT 240
Db 174064 TGTCTATAGAACAAATTTTATTTGTTCTGATATATGACTTTTATTTGTTATATGCTATACAACT 174005

QY 241 AGATTTTGTGTTTGTGTTTGTGACCGAGTCTTACTCTCTACCCAGGCTGGAGTGTAAATGGC 300
Db 174004 AGATTTTGTGTTTGTGTTTGTGACCAAGTCTTACTCTCTACCCAGGCTGGAGTGTAAATGGC 173945

QY 301 ATGCTCTAGCTCACTGCAACCTCCGCTCCCGGG 335
Db 173944 ATGCTCTAGCTCACTGCAACCTCCGCTCCCGAGG 173910

RESULT 11
AC074003
LOCUS 246865 bp DNA linear HTG 24-SEP-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-163G14, *** SEQUENCING IN
PROGRESS ***, 53 unordered pieces.
AC074003
AC074003.3 GI:10281455
VERSION HTG: HTGS PHASE1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 246865)
AUTHORS Mammaliya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Waterston,R.H.
JOURNAL The sequence of Homo sapiens clone
AUTHORS Unpublished
TITLE 2 (bases 1 to 246865)
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
TITLE Submitted (09-JUL-2000) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Sep 24, 2000 this sequence version replaced gi:9186850.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0163G14
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 213368 bases at least Q40
Consensus quality: 223527 bases at least Q30
Consensus quality: 229418 bases at least Q20
Insert size: 170000; agarose-fp
Quality coverage: 4.83 in Q20 bases; agarose-fp
Quality coverage: 3.37 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
*
* 1 12900: contig of 12900 bp in length
* 13001 13000: gap of unknown length
* 23006 23006: contig of 10006 bp in length
* 23106 23106: gap of unknown length
* 23107 35032: contig of 11926 bp in length
* 35132 35132: gap of unknown length
```

```

* 35133 50532: contig of 15400 bp in length
* 50533 50632: gap of unknown length
* 50633 64756: contig of 14124 bp in length
* 64757 64856: gap of unknown length
* 64857 80978: contig of 16122 bp in length
* 80979 81078: gap of unknown length
* 81079 101159: contig of 20081 bp in length
* 101160 10259: gap of unknown length
* 101260 102865: contig of 1606 bp in length
* 102866 102965: gap of unknown length
* 102966 104593: contig of 1628 bp in length
* 104594 104693: gap of unknown length
* 104694 106231: contig of 1338 bp in length
* 106232 106331: gap of unknown length
* 106332 107693: contig of 1362 bp in length
* 107694 107793: gap of unknown length
* 107794 109155: contig of 1362 bp in length
* 109156 109255: gap of unknown length
* 109256 110481: contig of 1226 bp in length
* 110482 110581: gap of unknown length
* 110582 111881: contig of 1300 bp in length
* 111882 111981: gap of unknown length
* 111982 113524: contig of 1543 bp in length
* 113525 113624: gap of unknown length
* 113625 114983: contig of 1359 bp in length
* 114984 115083: gap of unknown length
* 115084 116145: contig of 1062 bp in length
* 116146 116245: gap of unknown length
* 116246 117625: contig of 1380 bp in length
* 117626 117725: gap of unknown length
* 117726 120009: contig of 2284 bp in length
* 120010 120109: gap of unknown length
* 120110 122469: contig of 2360 bp in length
* 122470 122569: gap of unknown length
* 122570 124001: contig of 1432 bp in length
* 124002 124101: gap of unknown length
* 124102 126352: contig of 2251 bp in length
* 126353 126452: gap of unknown length
* 126453 128477: contig of 2025 bp in length
* 128478 128577: gap of unknown length
* 128578 130466: contig of 1889 bp in length
* 130467 130566: gap of unknown length
* 130567 132508: contig of 1939 bp in length
* 132509 132605: gap of unknown length
* 132606 134148: contig of 1543 bp in length
* 134149 134248: gap of unknown length
* 134249 136448: contig of 2200 bp in length
* 136449 136548: gap of unknown length
* 136549 138255: contig of 1707 bp in length
* 138256 138355: gap of unknown length
* 138356 140535: contig of 2180 bp in length
* 140536 140635: gap of unknown length
* 140636 142199: contig of 1564 bp in length
* 142200 142299: gap of unknown length
* 142300 144900: contig of 2607 bp in length
* 144901 145008: gap of unknown length
* 145009 147213: contig of 2207 bp in length
* 147214 147313: gap of unknown length
* 147314 148878: contig of 1565 bp in length
* 148879 148978: gap of unknown length
* 148979 150908: contig of 1928 bp in length
* 150909 151008: gap of unknown length
* 151010 153354: contig of 2348 bp in length
* 153355 153454: gap of unknown length
* 153455 155071: contig of 1617 bp in length
* 155072 155171: gap of unknown length
* 155172 159010: contig of 3839 bp in length
* 159011 159110: gap of unknown length
* 159111 162070: contig of 2960 bp in length
* 162071 162170: gap of unknown length
* 162171 164452: contig of 2282 bp in length
* 164453 164552: gap of unknown length
* 164553 166294: contig of 3742 bp in length

* 168295 168394: gap of unknown length
* 168395 171385: contig of 2991 bp in length
* 171386 171485: gap of unknown length
* 171486 175016: contig of 3331 bp in length
* 175017 175116: gap of unknown length
* 175117 178766: contig of 3650 bp in length
* 178767 178866: gap of unknown length
* 178867 182838: contig of 3972 bp in length
* 182839 182938: gap of unknown length
* 182939 186148: contig of 3210 bp in length
* 186149 186248: gap of unknown length
* 186249 190168: contig of 3920 bp in length
* 190169 190268: gap of unknown length
* 190269 193524: contig of 3256 bp in length
* 193525 193624: gap of unknown length
* 193625 202160: contig of 8536 bp in length
* 202161 202260: gap of unknown length
* 202261 210133: contig of 7873 bp in length
* 210134 210233: gap of unknown length
* 210234 218141: contig of 7908 bp in length
* 218142 218241: gap of unknown length
* 218242 228213: contig of 9972 bp in length
* 228214 228313: gap of unknown length
* 228314 237496: contig of 9183 bp in length
* 237497 237597: gap of unknown length
* 237597 246865: contig of 9269 bp in length.

FEATURES
Location/Qualifiers
1. 246865
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-163G14"
1. 12900
/note="assembly_name:Contig100"
13001..23006
/note="assembly_name:Contig101"
23107..35032
/note="assembly_name:Contig102"
35133..50532
/note="assembly_name:Contig103"
50633..64756
/note="assembly_name:Contig104"
64857..80978
/note="assembly_name:Contig105"
81079..101159
/note="assembly_name:Contig106"
101260..102865
/note="assembly_name:Contig107"
102966..104593
/note="assembly_name:Contig108"
104694..106231
/note="assembly_name:Contig109"
106332..107693
/note="assembly_name:Contig110"
107794..109155
/note="assembly_name:Contig111"
109256..110481
/note="assembly_name:Contig112"
110582..111881
/note="assembly_name:Contig113"
111882..113524
/note="assembly_name:Contig114"
113625..114983
/note="assembly_name:Contig115"
115084..116145
/note="assembly_name:Contig116"
116246..117625
/note="assembly_name:Contig117"
117726..120009
/note="assembly_name:Contig118"
120010..122469
/note="assembly_name:Contig119"
122470..122569
/note="assembly_name:Contig120"
122570..124001
/note="assembly_name:Contig121"
124002..124101
/note="assembly_name:Contig122"
124102..126352
/note="assembly_name:Contig123"
126353..126452
/note="assembly_name:Contig124"
126453..128477
/note="assembly_name:Contig125"
128478..128577
/note="assembly_name:Contig126"
128578..130466
/note="assembly_name:Contig127"
130467..130566
/note="assembly_name:Contig128"
130567..132508
/note="assembly_name:Contig129"
132509..132605
/note="assembly_name:Contig130"
132606..134148
/note="assembly_name:Contig131"
134149..134248
/note="assembly_name:Contig132"
134249..136448
/note="assembly_name:Contig133"
136449..136548
/note="assembly_name:Contig134"
136549..138255
/note="assembly_name:Contig135"
138256..138355
/note="assembly_name:Contig136"
138356..140535
/note="assembly_name:Contig137"
140536..140635
/note="assembly_name:Contig138"
140636..142199
/note="assembly_name:Contig139"
142200..142299
/note="assembly_name:Contig140"
142300..144900
/note="assembly_name:Contig141"
144901..145008
/note="assembly_name:Contig142"
145009..147213
/note="assembly_name:Contig143"
147214..147313
/note="assembly_name:Contig144"
147314..148878
/note="assembly_name:Contig145"
148879..148978
/note="assembly_name:Contig146"
148979..150908
/note="assembly_name:Contig147"
150909..151008
/note="assembly_name:Contig148"
151009..153354
/note="assembly_name:Contig149"
153355..153454
/note="assembly_name:Contig150"
153455..155071
/note="assembly_name:Contig151"
155072..155171
/note="assembly_name:Contig152"
155172..159010
/note="assembly_name:Contig153"
159011..159110
/note="assembly_name:Contig154"
159111..162070
/note="assembly_name:Contig155"
162071..162170
/note="assembly_name:Contig156"
162171..164452
/note="assembly_name:Contig157"
164453..164552
/note="assembly_name:Contig158"
164553..166294
/note="assembly_name:Contig159"

Query Match 66.5%; Score 222.8; DB 2; Length 246865;

```



```
misc_feature 60141..61238
              /note="assembly_fragment"
misc_feature 61339..65067
              /note="assembly_fragment"
misc_feature 65168..107152
              /note="assembly_fragment"
misc_feature 107253..154782
              /note="assembly_fragment"
misc_feature 154883..161940
              /note="assembly_fragment"
              clone end:T7
              vector_side:right"

ORIGIN
Query Match      29.9%; Score 100; DB 2; Length 161940;
Best Local Similarity 62.11%; Pred.No.1.2e-08;
Matches 175; Conservative 0; Mismatches 105; Indels 2; Gaps 1;

QY 54 TCTGTTTGGGAAATTTTAAACTACTGCTTTTAAATTTCTTTATTTGTTGCTGTAATGACTA 113
Db 91637 TTTATATATTTATATATGTTTATATATATTTATATATGTTTATATATGTTTATATATTTA 91696

QY 114 TTTTACGTCATATACAAATTTTATGTTTGTAAATGACTTTATTTGTTGTCATATGAT 173
Db 91697 TATATGTTTATATATATATTTATATATATGTTTATATATTTATATATATTTTATATTTT 91756

QY 174 AATTTTATGTCATAGACAAATTTTATGTTGTTGATATATGACTTTTATTTTATATGCGCTA 233
Db 91757 TATATTTTATGTTTATATATATTTATATATTTTATATATTTTATATATTTTATATTTA--TA 91814

QY 234 TACAACTAGATTTTCTGTTGTTTGGACCGAGTCTACTCTGTCACCCAGCTGGAGTG 293
Db 91815 TATATATATATTTTATTTTATTTTGGACCGAGTCTGCTCTGTCACCCAGCTGGAGTG 91874

QY 294 TAATGGCATGCTCAGCTCACTGCAACTCCGCTCCCGG 335
Db 91875 CAGTGGCAGCATCTGGCTCACTACAGCTCCACTCCCGG 91916

RESULT 13
AC073310 LOCUS 125685 bp DNA linear PRI 08-OCT-2003
DEFINITION Homo sapiens BAC clone RP11-49G5 from 7, complete sequence.
ACCESSION AC073310
VERSION AC073310.7 GI:13176604
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 125685)
AUTHORS Sulston,J.E. and Wilson,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 125685)
AUTHORS Haakenson,W., Hannah,C. and Kang,K.
TITLE The sequence of Homo sapiens BAC clone RP11-49G5
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 125685)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 125685)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 125685)

AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 125685)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 125685)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
8 (bases 1 to 125685)
AUTHORS Wilson,R.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 1, 2001 this sequence version replaced gi:11024926.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH00049G05
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/GRB/CHR7, send
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu

SOURCE INFORMATION:
The RP11-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Moon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-409M7, 200 bp overlap the
clone sequenced to the right is RP11-374N8. Actual start of this
clone is at base position 14572 of RP11-409M7 actual end is at base
position 125685 of RP11-49G5.
Location/Qualifiers
1. 125685
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"

FEATURES
Source
```

[illegible]

Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeAvellino, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Headford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., Menes, L., Minova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (24-AUG-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Dec 13, 2000 this sequence version replaced gi:9280765.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIDR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2203

Center clone name: 325_L_12

FEATURES

source

Location/Qualifiers

1. .184864
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RPL1-325L12"
/clone_lib="RPL1-11 Human Male BAC"

286. .542

/rpt_family="AluJo"

complement(1302. .1593)

/rpt_family="AluSc"

1769. .2080

/rpt_family="AluSx"

2099. .2161

/rpt_family="L2"

2215. .2322

/rpt_family="L2"

2673. .2700

/rpt_family="AT_rich"

2742. .2856

/rpt_family="MIR3"

2888. .3180

/rpt_family="AluSp"

3289. .3597

repeat_region
3936. .4107
/rpt_family="AluY"
repeat_region
4158. .4204
/rpt_family="FAM"
repeat_region
4379. .4492
/rpt_family="L2"
repeat_region
4500. .4543
/rpt_family="MIR"
repeat_region
4544. .4846
/rpt_family="MIR"
repeat_region
4847. .4972
/rpt_family="AluSx"
repeat_region
5225. .5333
/rpt_family="MIR"
repeat_region
5363. .5655
/rpt_family="MIR3"
repeat_region
5656. .5781
/rpt_family="LIMASa"
repeat_region
5790. .6207
/rpt_family="MLT2D"
repeat_region
6313. .6695
/rpt_family="LIMASa"
repeat_region
6696. .6998
/rpt_family="MLT2B1"
repeat_region
6999. .7108
/rpt_family="AluJb"
repeat_region
7117. .7273
/rpt_family="MLT2B1"
repeat_region
7485. .7608
/rpt_family="MIR"
repeat_region
7485. .7608
/rpt_family="MERSA"
repeat_region
8299. .8609
/rpt_family="AluSp"
repeat_region
9109. .9192
/rpt_family="MIR"
repeat_region
9302. .9643
/rpt_family="L1PA10"
repeat_region
9647. .9916
/rpt_family="L2"
repeat_region
10055. .10271
/rpt_family="L2"
repeat_region
10361. .10480
/rpt_family="MER94"
repeat_region
10715. .11025
/rpt_family="AluJo"
repeat_region
11563. .11563
/rpt_family="AluY"
repeat_region
11818. .12119
/rpt_family="AluSg"
repeat_region
12564. .12623
/rpt_family="MADE1"
repeat_region
12710. .13002
/rpt_family="AluSx"
repeat_region
12967. .12977
/rpt_family="AluSx"
repeat_region
13004. .13008
/note="<30 qual SNGL region"
repeat_region
13154. .13452
/rpt_family="AluSp"
repeat_region
14745. .15042
/rpt_family="AluSx"
repeat_region
15514. .15810
/rpt_family="AluY"
repeat_region
16087. .16465
/rpt_family="MER61D"
repeat_region
16517. .16762
/rpt_family="L1MC4a"
repeat_region
17245. .17342
/rpt_family="L1MB1"
repeat_region
17599. .17737
/rpt_family="MIR3"
repeat_region
18456. .18775
/rpt_family="L2"

Mon May 24 08:12:32 2004

```
repeat_region complement(18776..19087)
repeat_region /rpt family="AluJb"
19088..19313
repeat_region /rpt family="L2"
repeat_region complement(19342..20175)
20176..20477
repeat_region /rpt family="L1MB7"
repeat_region /rpt family="AluSx"
repeat_region complement(20478..20620)
repeat_region complement(20621..20931)
repeat_region /rpt family="AluJb"
repeat_region complement(20932..21437)
repeat_region /rpt family="L1MB7"
repeat_region 21438..21459
repeat_region /rpt family="(TAA)n"
repeat_region complement(21460..21545)
repeat_region /rpt family="L1MB7"
repeat_region complement(21546..21834)
repeat_region /rpt family="AluSx"
repeat_region complement(21835..22416)
repeat_region /rpt family="L1MB7"
repeat_region complement(22417..22548)
repeat_region /rpt family="FLAM A"
repeat_region complement(22549..22796)
repeat_region /rpt family="L1MB7"
repeat_region 22797..23107
repeat_region /rpt family="AluSx"
repeat_region complement(23108..23247)
repeat_region /rpt family="L1MB7"
repeat_region 23255..23457

Query Match 28.6%; Score 95.8; DB 9; Length 184864;
Best Local Similarity 57.3%; Pred. No. 6e-08;
Matches 192; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

QY 2 ATTATTTTTCCTTAACTTCAAACTCAAGGAAACCCAGTTGGCGCTGACTCTGTTG 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155318 ATTATCTTAAATTTTAAATTTTAAATGTCCTCAATGAAATGCAATATTTTATGTTTATA 155259

QY 62 TGGAAATTTTAACTACTGTTTAAATTTCTTTATTTGTTGTTATATGCTATTTTACGT 121
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155258 TATATATTTTATATATATTTTATATATATTTTATATATATTTTATATATATTTTATA 155199

QY 122 CATATAACAATTTTATTTGTTTAAATGACTTATTTGTTGTCATATGATAATTTTAT 181
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155198 TATATATTTTATATATATTTTATATATATTTTATATATATTTTATATATATTTTATA 155139

QY 182 GTCATAGAACAAATTTTATGCTTCATATATGACTTTAT-TGTTATATGGCTATACAACT 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155138 TATATATTTTATATATATTTTATATATATTTTATATATATTTTATATATATTTTAT 155079

QY 241 AGATTTTGTGTTTGTGTTTGTGCGAGTTTACTCTGTCACCCAGCTGAGTGAATGCG 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155078 ATATTTTGTGTTTGTGTTGAGACGGAGTCTGCTCTGTTGCCAGCTGGAGTGCAGTGGC 155019

QY 301 ATGGTCTCAGCTCACTGCAACTCCGCTCCCGGG 335
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155018 CGATCTCAGCTCACTGCAAGCTCCACCTCCAGG 154984
```

```
RESULT 15
AC073251
LOCUS AC073251 163401 bp DNA linear HTG 07-JUL-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-102D20, WORKING DRAFT
SEQUENCE, 33 unordered pieces.
ACCESSION AC073251
VERSION AC073251.2 GI:8954214
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 163401)
Waterson,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 163401)
Waterson,R.H.
Direct Submission
Submitted (12-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 7, 2000 this sequence version replaced gi:8468953.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0102D20
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: Plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 145004 bases at least Q40
Consensus quality: 150190 bases at least Q30
Consensus quality: 153173 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 160201; sum-of-contigs
Quality coverage: 3.36 in Q20 bases; agarose-fp
Quality coverage: 3.59 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1128: contig of 1128 bp in length
* 1129: gap of unknown length
* 1229: contig of 1293 bp in length
* 2522: gap of unknown length
* 2622: contig of 1359 bp in length
* 3981: gap of unknown length
* 4081: contig of 2197 bp in length
* 6278: gap of unknown length
* 6378: contig of 2122 bp in length
* 8500: gap of unknown length
* 8600: contig of 2287 bp in length
* 10887: gap of unknown length
* 10986: contig of 2748 bp in length
* 13735: gap of unknown length
* 13835: contig of 3140 bp in length
* 16975: gap of unknown length
* 17075: contig of 3713 bp in length
* 20788: gap of unknown length
* 20888: contig of 2504 bp in length
* 23392: gap of unknown length
* 23492: contig of 3757 bp in length
* 27249: gap of unknown length
* 27349: contig of 3400 bp in length
* 30849: gap of unknown length
* 34129: contig of 3280 bp in length
* 34229: gap of unknown length
* 37276: contig of 3047 bp in length
* 37376: gap of unknown length
* 40709: contig of 3333 bp in length
* 40809: gap of unknown length
* 44045: contig of 3236 bp in length
* 44145: gap of unknown length
* 47939: contig of 3795 bp in length

